1989

The comparison of diagnostic tests when one is based upon a discrete separator variable and the others are based upon continuous separator variables

Craig Allen Beam
Iowa State University

Order Number 8920112

The comparison of diagnostic tests when one is based upon a discrete separator variable and the others are based upon continuous separator variables

Beam, Craig Allen, Ph.D.

Iowa State University, 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

The comparison of diagnostic tests when one is based upon a discrete separator variable and the others are based upon continuous separator variables

by

Craig Allen Beam

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1989

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Diagnostic Tests

A diagnostic test is a rule for classifying an individual as belonging to one of two populations. For convenience, we will call one population the Negative Population and the other the Positive Population.[1] This classification rule is based upon the realized value of a Separator Variable, a scalar-valued measurement made upon the individual, and a Positivity Region. The Positivity Region is a sub-set of the domain of the Separator Variable which, when observed, causes the individual to be classified as belonging to the Positive Population. When the value of the Separator Variable falls outside of this Region, the individual is classified as belonging to the Negative Population.

For purposes of simplicity, my dissertation will, with the exception of a special case, be concerned only with diagnostic tests based upon real-valued separator variables whose Positivity Region can be specified by an interval of the form $\tilde{P} = (c,\infty)$. That is, values of the Separator Variable greater than some number c cause the individual to be classified as belonging to the Positive Population. To the best of my knowledge, diagnostic tests of this form are exclusively considered in the Literature: Not only are diagnostic tests based upon real-valued measurements, but also a commonly found assumption is that "large values" of the Separator Variable are "associated" with the Positive Population. In this context,

---

[1] Some of the terminology associated with diagnostic tests has not yet been standardized. Precedence for the terms and definitions used in this section can be found in Weinstein and Fineberg (1980).

c is herein referred to as the Positivity Criterion or, equivalently, Cutoff Value for the diagnostic test.

Now, letting D^- denote the Negative Population and D^+ the Positive Population and τ denote a diagnostic test based upon the Separator Variable ξ, we may summarize the previous ideas by writting

$$\tau = \begin{cases} T^- \text{ (classify as belonging to } D^-) \iff \xi \leq c \\ T^+ \text{ (classify as belonging to } D^+) \iff \xi > c \end{cases}$$

where $c \in \mathbb{R}^1$ denotes a chosen cutoff value.

1.2 Probabilistic Attributes of Diagnostic Tests

We now consider the following experiment: An individual will be selected at random from a population composed of a mixture of individuals from D^- and D^+ and is to be classified, i.e., "diagnosed," via the diagnostic test τ. The outcomes of the experiment, and their names as usually appear in the Literature, can be displayed by the outcome matrix displayed in Figure 1.

		True Membership	
		D^-	D^+
Test Result	T^-	True Negative (TN)	False Negative (FN)
	T^+	False Positive (FP)	True Positive (TP)

Figure 1. Outcomes and their nomenclature from the experiment of diagnosing a randomly selected individual

Let $P(D^-)$ be the probability the individual selected is a member of the Negative Population and $P(D^+)$ be the probability the individual is a member of the Positive Population. Then, with an obvious extension of notation, probabilities associated with the various outcomes may be identified to be

$$P(TN) = P(D^-)P(T^-/D^-)$$

$$P(FN) = P(D^+)P(T^-/D^+)$$

$$P(FP) = P(D^-)P(T^+/D^-)$$

$$P(TP) = P(D^+)P(T^+/D^+).$$

When thinking about D^+ as the population of "diseased" individuals, standard terminology for $P(D^+)$ is "Prevalence Rate." In the context of medical decision making based upon diagnostic tests, $P(D^+)$ may be referred to as the "Prior Probability of Disease" - prior here referring to before information from the test is obtained (Weinstein and Fineberg (1980)).

Now, if one were to classify an individual soley on the basis of knowledge of the prevalence rate in the population one might classify the individual as belonging to D^- with probability $P(D^-)$ and classify as belonging to D^+ with probability $P(D^+)$ and then the outcome probabilities would be

$$P(TN) = P(D^-)P(D^-)$$

$$P(FN) = P(D^+)P(D^-)$$

$$P(FP) = P(D^-)P(D^+)$$

$$P(TP) = P(D^+)P(D^+).$$

If a diagnostic test is to be useful it ought to improve the chances of making a correct diagnosis over that afforded simply by knowledge of prevalence rates in the population. Since improvement in the diagnostic procedure will be realized if, and only if, $P(T^-/D^-) > P(D^-)$ and $P(T^+/D^+) > P(D^+)$, it stands to reason that pertinent attributes of a diagnostic test are these conditional probabilities of making a correct diagnosis.

In light of this, much of the work in the assessment of diagnostic tests has concentrated upon these two probabilistic attributes and standard terminologies for them have evolved. We therefore now define the Specificity of a test to be the probability the test correctly classifies an individual from D^-, i.e., $P(T^-/D^-)$, and the Sensitivity[2] of a test to be the probability the test correctly classifies an individual from D^+, i.e., $P(T^+/D^+)$.

It is worthwhile to point out that our discussion of the assessment of diagnostic tests is, at this juncture, completely analogous to the usual theoretical considerations given to the assessment of a statistical hypothesis test of a simple null versus a simple alternative. For example, we are essentially concerned with the properties of a test of H_0: D^- versus H_A: D^+ based upon the observed value of a random variable (the Separator Variable). Historically, the statistically important aspects of this problem have been the size of the hypothesis test (which is equivalent to 1- the Specificity of the diagnostic test) and the power of the hypothesis test (equivalent to Sensitivity of the diagnostic test). Thus,

---

[2]Yerushalmy (1947) is credited with introduction of these two terms.

interest in the Specificity and Sensitivity of a diagnostic test may be seen to be a natural extension of interest in the size and power of a hypothesis test.

Other probabilistic attributes of diagnostic tests have been considered as well. Some researchers, for example, have been concerned with the inverse of the above probabilities: "Given the diagnostic test is positive (i.e., classifies the individual as belonging to D^+) what is the probability the individual does indeed belong to the Positive Population?" (An attribute known as the Predictive Value of a Positive Result.) Other researchers have addressed the overall performance of the test in what has come to be known as the Accuracy of the test which is the probability the test makes a correct diagnosis.[3]

Noting that the Predictive Value of a Positive Result

$$= P(D^+/T^+) = \frac{P(D^+)P(T^+/D^+)}{P(D^+)P(T^+/D^+) + P(D^-)P(T^+/D^-)}$$

and that

$$\text{Accuracy} = P(D^+)P(T^+/D^+) + P(D^-)P(T^-/D^-)$$

we see that these attributes are, however, functions of the prevalence rate in the population (which Specificity and Sensitivity are not) and are themselves functions of Specificity and Sensitivity.

---

[3] A discussion of various test attributes as well as an overview of this entire area appears in Begg (1986).

For the above reasons, Specificity and Sensitivity have been commonly used for the assessment and comparison of diagnostic tests. My work continues in this fashion, examining how to compare diagnostic tests on the basis of Specificity and Sensitivity.

1.3 The Comparison of Diagnostic Tests

Given a choice of several diagnostic tests and wanting to choose from among them a "best" test one would naturally select the test which imparts the greatest diagnostic "information." In light of the preceding discussion, we might regard the amount of information conveyed by a diagnostic test to be measured by some combination of the probabilistic test attributes Specificity and Sensitivity.

In certain situations, exclusive attention may justifiably be focused on just one of these two attributes. For example, one might want to apply a diagnostic test to a large, unselected group of individuals in hopes of detecting a rare (i.e., low prevalence) but extremely deadly disease (e.g., AIDS). One might have several candidate diagnostic tests, or "diagnostic screens" as they would in this case be called, at hand to choose from and might be inclined to want the screen with the highest probability of detecting diseased individuals with little or no concern for the rate at which nondiseased individuals might be falsely classified.

In such cases, one would want to compare diagnostic tests on the basis of a single attribute for which, of course, a number of standard statistical methods exist. A review of the problem of comparing diagnostic tests with respect to a single probabilistic attribute (and not just

Specificity or Sensitivity) and an encapsulation of the problem in a categorical analytic framework may be found in White and Landis (1982).

Of course, diagnosticians are often concerned with both the Specificity and Sensitivity of the procedures to be compared and it only stands to reason that one would always want to choose the procedure having both greatest Specificity and greatest Sensitivity. However, such strict dominance of one test over another may just not be found. For example, it might be the case that although test A is more sensitive than test B (i.e., has greater Sensitivity), it however is not as specific as B (i.e., test B has greater Specificity).

One solution to the problem is to combine the two probabilistic attributes into a single measure. Youden (1950), for example, offered the combination

$$J = \text{Sensitivity} + \text{Specificity} - 1.$$

Of course, since different combinations of Specificity and Sensitivity could yield the same value of J, this approach suffers a nonidentifiability deficit.

Two other solutions arise by imposing the restriction that the diagnostic tests be based upon separator variables which are real-valued measurements having continuous Cumulative Distribution Functions (cdfs) as described below and that large values of the Separator Variables are associated with membership in the positive population.

Specifically, suppose we have two diagnostic tests, τ_1 and τ_2,

to compare. Let ξ_i represent the separator variable for test τ_i, $\xi_i \in \mathbb{R}^1$. We envision there to be a distribution of the values of ξ_i when measured upon individuals from D^- and a distribution when measured upon individuals from D^+. We shall call these distributions F_i and G_i, respectively, and require that they be continuous functions of ξ_i.

As discussed earlier, we may now describe the tests succinctly as

$$\tau_i = \begin{cases} T_i^- & \text{if} \quad \xi_i \leq c_i \\ T_i^+ & \text{if} \quad \xi_i > c_i \end{cases}, \quad i = 1,2,$$

and, letting $S_p(c_i)$ = Specificity of test τ_i and $S_e(c_i)$ = Sensitivity of test τ_i, it then follows immediately that

$$S_p(c_i) = F_i(c_i)$$

and (1.3.1)

$$S_e(c_i) = 1 - G_i(c_i) ,$$

and that these probabilistic attributes are functions of the cutoff values c_i.

One solution to the problem of comparing diagnostic tests when both Specificity and Sensitivity are of interest is given by Greenhouse and Mantel (1950). Here, the authors contend that the only way to fairly compare two diagnostic tests with respect to, say, their Sensitivity is to require them to first have common Specificity.

Following this, Greenhouse and Mantel (1950) propose the following method for the comparison of the Sensitivities of two diagnostic tests when they are both based upon continuous separator variables:

a) Select a, perhaps clinically relevant, Specificity p.

b) Choose the cutoff for each test so that both possess Specificity p.

c) Then compare the Sensitivities of the test by taking a difference.

Symbolically, Greenhouse and Mantel (1950) thus identify as the parameter of interest in comparing two diagnostic tests with respect to both Specificity and Sensitivity to be

$$[1 - G_1F_1^{-1}(p)] - [1 - G_2F_2^{-1}(p)].$$

Further evaluation of the Greenhouse-Mantel method can be found in Linnet (1987).

The other solution to the comparison of diagnostic tests on the basis of Specificity and Sensitivity compares the tests not just at a single specificity-sensitivity combination but at all such combinations available to each test in the following manner.

By (1.3.1), choice of cutoff dictates the Specificity and Sensitivity of the test. Letting c_i range across the real numbers we obtain the entire range of specificity-sensitivity combinations available to test τ_i. Such combinations may be represented via a device known as a Receiver Operating Characteristic, or ROC, curve which, usually, is a

plot of $S_e(c_i)$ versus $1-S_p(c_i)$. A typical ROC curve is represented in Figure 2.

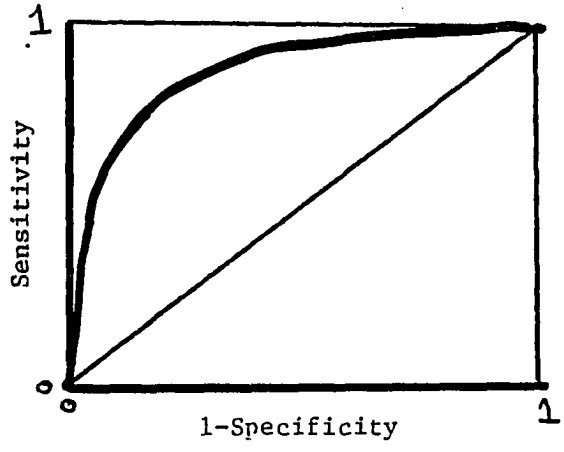

Figure 2. A typical ROC curve

The 45^o line in Figure 2 represents Specificity-Sensitivity combinations available to a test that is based purely upon a randomization device. A diagnostic test ought to do better than this, and so the ROC curve typically arches above the 45^o line. Since we consider greater Specificity and greater Sensitivity to be associated with "better" tests, we conclude that the greater the arch in the ROC curve the better the test.

One measure proposed to capture the entire range of Specificity-Sensitivity combinations available to a test is the area under the ROC curve (see, for an example, Green and Swets (1966) for further discussion)

and by the above discussion it follows that if one test is consistently "better" than another, the area under its ROC curve will be greater than the other. Thus, another way of comparing two diagnostic tests which are based upon continuous separator variables is to compare the areas under their ROC curves.

One method of estimating the area under an ROC curve is to use maximum likelihood estimation when the distributions F and G are normal (see Dorfman and Alf (1969)). Another method is to use the trapezoidal rule which, in this case, is equivalent to the Wilcoxon statistic. In fact, the area itself is the probability that the value of the Separator Variable when measured on a randomly selected member from D^+ exceeds the value when measured on a randomly selected member from D^- (see Bamber (1975)). Appropriate estimation methods for independent versus dependent sampling and grouped versus ungrouped data have been developed (see, for example, Hanley and McNeil (1982, 1983); Delong et al. (1988)).

Of course, a problem with the use of the area under the ROC curve to compare tests is that although the total areas from two are equal, one ROC curve might yet dominate the other over a region of clinically desirable specificities. A solution to this problem is given by Wieand et al. (1987) who provide a general framework for the comparison of the areas under two ROC curves over restricted regions of Specificity.

1.4 Comparing Diagnostic Tests When One is Based Upon a Discrete Separator Variable and the Other is Based Upon a Continuous Separator Variable

A problem yet to be considered, and one which has not been considered explicitly in the statistical literature, is how to compare two diagnostic tests when the separator variable of one test is a discrete variable and the separator of the other is continuous. This problem has arisen in the course of medical research. For example, Ahlquist et al. (1985) compare Hemoccult and HemoQuant, indicants of elevated fecal hemoglobin, as possible screens for gastrointestinal malignancies. While the HemoQuant assay measures miligrams of fecal hemoglobin per gram of stool, the Hemoccult test results in either a negative or positive reading and thus is a discrete separator.

Another example is given by Delong et al. (1988) who compared a discrete preoperative scoring system for use as a screening test in determining a patient's risk for failing to benefit from surgery versus two other preoperatively measured continuous indices. Here, the comparison is thus between a test based upon a discrete Separator Variable having thirteen possible values and two other tests each of which are based upon continuous measurements.

The immediately preceding comparison problem is used by the authors to demonstrate application of their statistical methodology for comparing the areas under "correlated" ROC curves. Their paper, however, does not address two important issues which place in doubt the appropriateness of ROC methodologies when one of the tests compared is not based upon a continuous Separator Variable.

The first issue here, as is demonstrated in the Ahlquist et al. (1985) study, is that the discrete Separator Variable need not be ordinal. Thus, in some cases an ROC curve as presently defined might not exist for one of the tests.

The second issue is that in many cases of practical interest the comparison of the areas under ROC curves may be severely biased against the discrete test simply by the way the curve is defined for these tests.

Consider, for example, the situation portrayed in Figure 3. Here, the ROC curve associated with the continuous test is the smooth concave curve. The series of line segments represent the ROC curve associated with the discrete test and shows how the ROC curve would be defined in this case by use of the Trapezoidal rule.

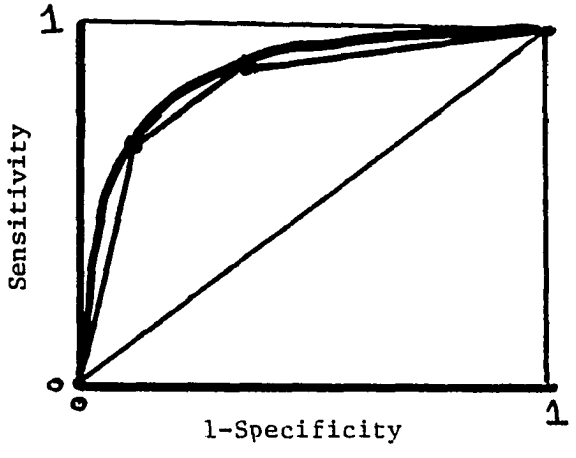

Figure 3. ROC curves of tests based upon discrete and continuous Separator Variables

Note that although the Sensitivities of the two tests are equal at each of the Specificities available to the discrete test, mere

concavity of the ROC curve for the continuous test ensures that it dominates the ROC curve of the discrete test. Thus, comparison of areas under these curves is biased against the discrete test. As concavity in the "continuous" ROC curve is not an unlikely occurrence, this example shows the ROC methodology to be inappropriate for a possibly widely occurring class of problems.[4]

Returning to the Ahlquist et al. (1985) paper, the Greenhouse-Mantel approach is not directly appropriate here either. For the Greenhouse-Mantel procedure requires the selection of a Specificity at which the Sensitivities of the tests are to be compared. But, the set of such Specificities now is constrained to the discrete set available to the test based upon the discrete separator. Furthermore, since these Specificities are not usually known, they must be estimated and this introduces a component of variability not considered in the work of Greenhouse and Mantel (1950).

My dissertation develops a methodology for the comparison of several diagnostic tests when one is based upon a discrete Separator Variable and the others are based upon continuous separator variables which is not afflicted by the above problems. In Chapter 2, my work extends the Greenhouse-Mantel statistic first to the situation where the discrete test is binary and then to where the discrete test is S-nary. Two diagnostic tests may therefore be compared by a linear combination of the

---

[4]Concavity will occur whenever the distribution of the continuous variable possesses the monotone likelihood ratio property.

Sensitivities of the tests when both are set to have the various Specificities estimated to be available to the discrete test. Thus, the procedure so developed exchanges the comparison of overall test performance (via the ROC) with a series of point-wise comparisons between the tests. Finally, the method is extended to the comparison of several tests based upon continuous Separators and one based upon a discrete Separator variable.

Chapter 3 reports the results of several Monte Carlo studies of the small-sample properties of the statistics developed in Chapter 2. Finally, Chapter 4 demonstrates application of the statistic to the data analyzed by Ahlquist et al. (1985) and by Delong et al. (1988).

2. SEVERAL STATISTICS AND THEIR ASYMPTOTIC DISTRIBUTION

2.1 General Considerations

For simplicity in presentation we will, in accordance with the standard applications of diagnostic tests, now refer to the Negative Population as the population of "controls" (e.g., nondiseased individuals) and to the Positive Population as the population of "cases" (e.g., diseased individuals). We assume that independent random samples of m controls and n cases may be selected.

We are to compare the T diagnostic tests, $\tau_1,\tau_2,\ldots,\tau_T$, of which τ_1 is based upon a discrete separator variable and the remainder are based upon continuous separators. For each of the m randomly selected controls we observe a value of the vector $\underline{X}_i' = (X_{1i},X_{2i},\ldots,X_{Ti})$, i = 1(1)m, where X_{1i} is the value of the discrete separator and $X_{2i},\ldots,X_{Ti}$ are the values of the continuous separators when measured upon the i-th control. Similarly, for each of the n randomly selected cases we observe a value of the vector $\underline{Y}_j' = (Y_{1j},Y_{2j},\ldots,Y_{Tj})$, j = 1(1)n, giving results from the discrete and the (T-1) continuous measurements, respectively.

Let S be the number of possible values for the discrete separator variable. When S > 2 we shall require the separator to be ordinal with large-ordered values being associated with cases. When S = 2 we shall not require ordinality, but for simplicity shall assume that the binary values have been coded as 0 or 1 with a value of 1 causing the individual to be classified as a "control" and a value of 0 to be classified as a

"case."[5]

We extend the procedure introduced by Greenhouse and Mantel (1950) in the following manner. In Section 2.2, we consider the case when T = 2 and S = 2. Letting F_2 and G_2 be the unknown cdfs of the continuous separator variable when applied to, respectively, controls and cases, and letting $p = \Pr\{X_{1i} = 1\}$, $i = 1(1)m$, $q = \Pr\{Y_{1j} = 1\}$, $j = 1(1)n$, where p and q are also unknown, and following the Greenhouse-Mantel method, the parameter for comparison is now seen to be

$$\Delta = [1-q] - [1 - G_2F_2^{-1}(p)] = G_2F_2^{-1}(p) - q.$$

Inserting the "usual" sample estimators of the components of the parameter; viz., estimating probabilities by sample proportions and estimating the inverse of a cdf by an order statistic, we identify the statistic to be

$$\hat{\Delta} = \widehat{G_2F_2^{-1}(p)-q} = \hat{G}_2\hat{F}_2^{-1}(\hat{p})-\hat{q}$$

$$= \frac{1}{n}\sum_{j=1}^{n} I(X_2(r;m)-Y_{2j}) - \frac{1}{n}\sum_{j=1}^{n} Y_{1j} ,$$

where

$$I(Z) = 1 \quad \text{if} \quad Z \geq 0$$

$$\qquad 0 \quad \text{if} \quad Z < 0$$

---

[5]Although apparently contrary to associating large values with cases, this assumption eases the theoretical development for the special case of a binary separator variable.

and

$$r = \sum_{i=1}^{m} X_{1i}$$

and

$X_2(r;m)$ is the r-th ordered value of $X_{21}, X_{22}, \ldots, X_{2m}$.

Thus, the extended procedure is to first estimate the specificity for the first test and then estimate the associated percentile of the distribution of results when the separator variable of the second test is applied to the controls. This estimated percentile is then treated as a cutoff point for the classification of individuals when using the second test. The sensitivity of the second test is then estimated using this cutoff and is compared, by taking a difference, with the estimated sensitivity of the first test.

Note that if the first test was based upon a continuous separator variable, the Greenhouse-Mantel statistic would be applicable and would have form

$$\hat{\Delta}_{GM} = \hat{G}_2\hat{F}_2^{-1}(p) - \hat{G}_1\hat{F}_1^{-1}(p)$$

$$= \frac{1}{n}\sum_{j=1}^{n} I(X_2(r^*;m)-Y_{2j}) - \frac{1}{n}\sum_{j=1}^{n} I(X_1(r^*;m)-Y_{1j})$$

where p is the specificity chosen for the comparison and $r^* = pm$.

The difference between the Greenhouse-Mantel statistic and its extension are immediate. First, when the first test is based upon a discrete separator the specificity p must be estimated and thus the order r is a random variable in the extended procedure whereas it is a preselected constant in the "nonextended" procedure. Second, the subtrahend of the extended procedure is not a function of an order statistic whereas the subtrahend of the Greenhouse-Mantel statistic is.

Theorem 1 of Section 2.2 establishes the asymptotic normality of the extended statistic under a few assumptions regarding continuity and differentiability of the cdfs involved. Theorem 1 is preceded by Lemma 1 which establishes the asymptotic normality of an order statistic when the order is a binomial random variable.

In Section 2.3, the Greenhouse-Mantel statistic is further extended to the case when $T \geq 2$ and $S \geq 2$. The statistic now becomes

$$\underline{\hat{\Delta}}_{(SxT)} =$$

$$\begin{bmatrix} \hat{G}_2\hat{F}_2^{-1}(\hat{p}_1)-\hat{q}_1 & \hat{G}_3\hat{F}_3^{-1}(\hat{p}_1)-\hat{q}_1 & \cdots & \hat{G}_T\hat{F}_T^{-1}(\hat{p}_1)-\hat{q}_1 \\ \hat{G}_2\hat{F}_2^{-1}(\hat{p}_2)-\hat{q}_2 & \hat{G}_3\hat{F}_3^{-1}(\hat{p}_2)-\hat{q}_2 & \cdots & \hat{G}_T\hat{F}_T^{-1}(\hat{p}_2)-\hat{q}_2 \\ \vdots & \vdots & & \vdots \\ \hat{G}_2\hat{F}_2^{-1}(\hat{p}_S)-\hat{q}_S & \hat{G}_3\hat{F}_3^{-1}(\hat{p}_S)-\hat{q}_S & \cdots & \hat{G}_T\hat{F}_T^{-1}(\hat{p}_S)-\hat{q}_S \end{bmatrix}$$

where

$$\hat{p}_s = P(X_{1i} \leq v_s)$$

$$\hat{q}_s = P(Y_{1j} \leq v_s)$$

and $v_1 < v_2 < \ldots < v_s$ are the S values attainable by the discrete separator variable.

Using certain assumptions, asymptotic normality of $\hat{\underline{\Delta}}$ is established by Theorem 2. This theorem utilizes Lemma 2 which is an extension of Lemma 1 to the multivariate distribution of S order statistics based upon orders r_s, s = 1(1)S, each of which has a binomial distribution. Lemma 2 shows this multivariate distribution to be asymptotically normal.

The asymptotic results of this chapter thus lay the foundation for the development of statistical methods for the comparison of diagnostic tests. These methods are developed in Chapter 3.

2.2 Comparing a Test Based Upon a Binary Separator with One Based Upon a Continuous Separator Variable

Lemma 1: Let (X_{11},X_{21}), (X_{12},X_{22}), ..., (X_{1m},X_{2m}) be a random sample from the bivariate distribution dF. Let the marginal distribution of X_1 be Bernoulli with $Pr\{X_{1i} = 1\} = p$, i = 1(1)m, and let the marginal cumulative distribution function of X_2 be denoted by $F_2(\cdot)$. Assume that F_2 is a continuous function, possessing finite derivatives of up to at least third order. Let $F(1,\xi) \equiv \int_{\{X_1=1, X_2 \leq \xi\}} dF$ and assume this

is a continuous function of ξ and that $F(1,F_2^{-1}(p)) \neq p$.

Define

$$r = \sum_{i=1}^{m} X_{1i}$$

$$\hat{p} = r/m$$

and $\hat{F}_2^{-1}(\hat{p}) = X_2(r;m)$ = the r-th ordered value of $\{X_{21},X_{22},\ldots,X_{2m}\}$ where we shall define $X_2(0;m) = -\infty$ and $X_2(m+1,m) = \infty$. Then,

$$\sqrt{m}[\hat{F}_2^{-1}(\hat{p})-F_2^{-1}(p)] \xrightarrow[m\to\infty]{L} N[0, \frac{2p-2F(1,F_2^{-1}(p))}{[f_2(F_2^{-1}(p))]^2}] .$$

Proof: Let c and k be real-valued constants. Note that $\hat{F}_2^{-1}(\hat{p}) \leq c$ <=> $X_2(r;m) \leq c$ <=> at least (r) of the X_{2i}'s $\leq c$ <=> (the proportion of X_{2i}'s $\leq c$) $\geq \frac{r}{m}$ <=> $\hat{F}_2(c) \geq \hat{p}$. And therefore,

$$P\{\sqrt{m}(\hat{F}_2^{-1}(\hat{p})-F_2^{-1}(p)) \leq k\} = P\{\hat{F}_2^{-1}(\hat{p}) \leq \frac{k}{\sqrt{m}} + F_2^{-1}(p)\}$$

$$= P\{\hat{F}_2(\frac{k}{\sqrt{m}} + F_2^{-1}(p)) \geq \hat{p}\}$$

$$= P\{\hat{F}_2(\frac{k}{\sqrt{m}} + F_2^{-1}(p)) - \hat{p} \geq 0\}$$

$$= P\{T_m^{(k)} \geq 0\}, \text{ say.}$$

We now show that $T_m^{(k)}$ has a limiting normal distribution. Let $b = (\frac{k}{\sqrt{m}} + F_2^{-1}(p))$, then we have,

$$T_m^{(k)} = \hat{F}_2(b) - \hat{p} = \frac{1}{m} \sum_{i=1}^{m} I(b-X_{2i}) - \frac{1}{m} \sum_{i=1}^{m} X_{1i}$$

$$= \frac{1}{m} \sum_{i=1}^{m} [I(b-X_{2i}) - X_{1i}] = \frac{1}{m} \sum_{i=1}^{m} \nu_{m_i}^{(k)},$$

where

$$\nu_{m_i}^{(k)} \equiv I(\frac{k}{\sqrt{m}} + F_2^{-1}(p) - X_{2i}) - X_{1i}.$$

Since the $\{(X_{1i},X_{2i})\}$, $i = 1,2,\ldots,m$ are iid it immediately follows that for each fixed value of m, the $\{\nu_{m_i}^{(k)}\}$ are iid. Also,

$$\mu_m^{(k)} \equiv E\nu_{m_i}^{(k)}, \quad i = 1,2,\ldots,m,$$

$$= E\{I(b-X_{2i}) - X_{1i}\} = F_2(b) - F_1(c).$$

$$\operatorname{cov}\{I(b-X_{2i}),X_{1i}\}$$

$$= E\{I(b-X_{2i})X_{1i}\} - E\{I(b-X_{2i})\}E\{X_{1i}\}$$

$$= F(1,b) - pF_2(b),$$

and therefore for fixed m,

$$\text{var}(\nu_m^{(k)}) \equiv \text{var}(\nu_{m_i}^{(k)}),\ i = 1,2,\ldots,m,$$

$$= \text{var}\{I(b-X_{2i})\} + \text{var}\{X_{1i}\} - 2\text{cov}\{I(b-X_{2i}),X_{1i}\}$$

$$= F_2(b)[1-F_2(b)] + p[1-p] - 2\{F(1,b)-pF_2(b)\}.$$

Let $\sigma_m^{(k)} \equiv \sqrt{\sum_{i=1}^{m} \text{var}(\nu_{m_i})}$. By identicality and independence when m is fixed, we have that

$$\sigma_m^{(k)} = \sqrt{m\text{var}(\nu_m)}. \tag{2.2.1}$$

Let $\omega_{m_i}^{(k)} = (\nu_{m_i}^{(k)} - \mu_m^{(k)})/\sigma_m^{(k)}$. It follows immediately that, for each m, the set of random variables, $\{\omega_{m_i}^{(k)}\}$, i = 1,2,...,m, are independent. Noting that, for each m and for all i = 1(1)m

$$\text{var}(\omega_{m_i}^{(k)}) = \frac{1}{m\text{var}(\nu_m)} \cdot \text{var}(\nu_m) = \frac{1}{m},$$

we have that, $\forall$ m and $\forall$ i, $\text{var}(\omega_{m_i}^{(k)}) < \infty$.

Following the Lindeberg-Feller Theorem, (see, for example, Chung (1974)), we may therefore establish that $\sum_{i=1}^{m} \omega_{m_i}^{(k)} \longrightarrow N(0,1)$ by showing that, for each $\eta > 0$,

$$\lim_{m\to\infty} \sum_{i=1}^{m} \int_{\{|\omega^{(k)}|>\eta\}} \omega^{(k)^2} dF_{\omega_{m_i}^{(k)}}(\omega) = 0, \tag{2.2.2}$$

where $F_{\omega_{m_i}^{(k)}}(\cdot)$ is the cdf of $\omega_{m_i}^{(k)}$.

Recalling the definition of $\omega_{m_i}^{(k)}$ and dropping the (k)-notation for the time being, we have that the L.H.S. of the above may be reexpressed as

$$\lim_{m\to\infty} \frac{1}{\sigma_m^2} \sum_{i=1}^{m} \int \left\langle \left| \frac{\nu-\mu_m}{\sigma_m} \right| > \eta \right\rangle (\nu-\mu_m)^2 dF_{\nu_{m_i}}^{(\nu)}, \tag{2.2.3}$$

where $F_{\nu_{m_i}}(\cdot)$ is the cdf of ν_{m_i}.

By (2.2.1), and identicality when m is fixed, (2.2.3)

$$= \lim_{m\to\infty} \frac{1}{m\mathrm{var}(\nu_m)} \cdot m \left(\int \left\langle \left| \frac{\nu-\mu_m}{\sigma_m} \right| > \eta \right\rangle (\nu-\mu_m)^2 dF_{\nu_m}^{(\nu)} \right),$$

where, when m is fixed, $F_{\nu_m}(\cdot)$ is the common cdf of ν_{m_i}, i = 1,2,...,m. Thus, to establish the "Lindeberg-Condition" (2.2.2) it suffices to show

$$\lim_{m\to\infty} \frac{1}{\mathrm{var}(\nu_m)} \left(\int \left\langle \left| \frac{\nu-\mu_m}{\sigma_m} \right| > \eta \right\rangle (\nu-\mu_m)^2 dF_{\nu_m}^{(\nu)} \right) = 0.$$

Now, $\lim_{m\to\infty} b = \lim_{m\to\infty}(k/\sqrt{m} + F_2^{-1}(p)) = F_2^{-1}(p)$, and as the functions $F_2(\xi)$ and $F(1,\xi)$ are continuous, we have that $\lim_{m\to\infty} \mathrm{var}(\nu_m)$

$$= \lim_{m\to\infty}\{p[1-p] + F_2(k/\sqrt{m} + F_2^{-1}(p))[1-F_2(k/\sqrt{m} + F_2^{-1}(p))]$$

$$- 2[F(1, k/\sqrt{m} + F_2^{-1}(p)) - pF_2(k/\sqrt{m} + F_2^{-1}(p))]\}$$

$$= p[1-p] + F_2(F_2^{-1}(p))[1-F_2(F_2^{-1}(p))]$$

$$- 2[F(1,F_2^{-1}(p)) - pF_2(F_2^{-1}(p))]$$

$$= p[1-p] + p[1-p] - 2[F(1,F_2^{-1}(p))-p^2]$$

$$= 2p[1-p] - 2[F(1,F_2^{-1}(p))-p^2]$$

$$= 2p - 2F(1,F_2^{-1}p).$$

Therefore, $\lim_{m\to\infty} \frac{1}{\text{var}(\nu_m)} = \frac{1}{2p-2F(c,F_2^{-1}(p))}$ exists, and will be finite unless $p = F(1,F_2^{-1}p)$.

However, by assumption we have that this cannot hold and thus we conclude that

$$\lim_{m\to\infty} \frac{1}{\text{var}(\nu_{m_i})} < \infty .$$

Now consider

$$\lim_{m\to\infty} \int_{\left\{\left|\frac{\nu-\mu_m}{\sigma_m}\right|>\eta\right\}} (\nu-\mu_m)^2 dF_{\nu_m}^{(\nu)} .$$

$$(\nu-\mu_m)^2 = (\nu_{m_i}-\mu_m)^2, \quad \text{for some} \quad i = 1,(1),m,$$

$$= (I(b-X_{2i}) - X_{1i} - [F_2(b)-p])^2$$

$$= ([I(b-X_{2i})-F_2(b)] + [p-X_{1i}])^2$$

$$= (x+y)^2, \text{ say.}$$

Note that $-1 \leq x \leq 1$ and $-1 \leq y \leq 1$ so that
$(x+y)^2 = x^2 + y^2 + 2xy \leq (\max|x|)^2 + (\max|y|)^2 + 2\max|x|\max|y| = 1+1+2 = 4.$
Thus,

$$0 \leq \lim_{m\to\infty} \int\left\{\left|\frac{\nu-\mu_m}{\sigma_m}\right|>\eta\right\}(\nu-\mu_m)^2 dF_{\nu_m}^{(\nu)} \leq 4\lim_{m\to\infty} \int\left\{\left|\frac{\nu-\mu_m}{\sigma_m}\right|>\eta\right\} dF_{\nu_m}^{(\nu)}.$$

But,

$$\int\left\{\left|\frac{\nu-\mu_m}{\sigma_m}\right|>\eta\right\} dF_{\nu_m}^{(\nu)} = P(\left|\frac{\nu-\mu_m}{\sigma_m}\right|>\eta) \leq E\,\frac{\{(\nu-\mu_m)^2/\sigma_m^2\}}{\eta^2}, \text{ by Chebyshev's Inequality,}$$

$$= \frac{\text{var}(\nu_m)}{m\text{var}(\nu_m)\eta^2} = \frac{1}{m\eta^2} \xrightarrow[m\to\infty]{} 0, \quad \forall\ \eta > 0.$$

Thus, since both limits exist and are finite, we have, $\forall\ \eta > 0$,

$$\lim_{m\to\infty} \frac{1}{\text{var}(\nu_m)}\left(\int\left\{\left|\frac{\nu-\mu_m}{\sigma_m}\right|>\eta\right\}(\nu-\mu_m)^2 dF_{\nu_m}^{(\nu)}\right)$$

$$= \lim_{m\to\infty} \frac{1}{\mathrm{var}(\nu_m)} \cdot \lim_{m\to\infty} \int_{\left\{\left|\frac{\nu-\mu_m}{\sigma_m}\right|>\eta\right\}} (\nu-\mu_m)^2 dF_{\nu_m}^{(\nu)} = 0.$$

That is, we have established condition (2.2.2).

As such we may conclude that

$$\sum_{i=1}^{m} \omega_{m_i}^{(k)} \xrightarrow[m\to\infty]{L} N(0,1),$$

or equivalently that

$$\sum_{i=1} \left(\frac{\nu_{m_i}^{(k)} - \mu_m^{(k)}}{\sigma_m^{(k)}} \right) \xrightarrow[m\to\infty]{L} N(0,1).$$

But, this implies that the distribution of $T_m^{(k)} = \frac{1}{m} \sum_{i=1}^{m} \nu_{m_i}^{(k)}$ is asymptotically normal.

Note that, for fixed m,

$$ET_m^{(k)} = E(\nu_m^{(k)}) = F_2(b) - p$$

$$= F_2(k/\sqrt{m} + F_2^{-1}(p)) - p.$$

Letting $\sigma^2_{T_m^{(k)}}$ denote the variance of $T_m^{(k)}$, we also have that

$$\sigma^2_{T_m^{(k)}} = \frac{1}{m} \mathrm{var}(\nu_m^{(k)})$$

$$= \frac{1}{m}(p[1-p] + F_2(b)(1-F_2(b)] - 2[F(1,b)-pF_2(b)])$$

$$= \frac{1}{m}(p[1-p] + F_2(k/\sqrt{m} + F_2^{-1}(p))[1-F_2(k/\sqrt{m} + F_2^{-1}(p))]$$

$$- 2[F(c, k/\sqrt{m} + F_2^{-1}(p) - pF_2(k/\sqrt{m} + F_2^{-1}(p))]).$$

Let $U_m^{(k)} = \sqrt{m}[\hat{F}_2^{-1}(\hat{p})-F_2^{-1}(p)]$ and recall that we had initially established the identity, $P(U_m^{(k)}<k) = P(T_m^{(k)}>0)$, from which it follows that

$$\lim_{m\to\infty} P(U_m^{(k)}<k) = \lim_{m\to\infty} P(T_m^{(k)}>0)$$

$$= \lim_{m\to\infty} P\left(\frac{T_m^{(k)}-E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}} > \frac{-E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}}\right)$$

$$= 1 - \lim_{m\to\infty} P\left(\frac{T_m^{(k)}-E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}} \leq \frac{-E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}}\right).$$

Consider the ratio

$$\frac{E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}} = \frac{F_2(k/\sqrt{m} + F_2^{-1}(p)) - p}{1/\sqrt{m}\ \sqrt{var(\nu_m)}}.$$

Expanding $E\{T_m^{(k)}\}$ about $F_2^{-1}(p)$ in a Taylor series gives

$$E\{T_m^{(k)}\} = F_2(F_2^{-1}(p)) + f_2(F_2^{-1}(p))(k/\sqrt{m})$$

$$+ f_2'(F_2^{-1}(p))(k/\sqrt{m})^2 + O[(k/\sqrt{m})^2] - p,$$

(where $f_2 = F_2'$),

$$= k/\sqrt{m}\ f_2(F_2^{-1}(p)) + O(1/m).$$

Thus,

$$\frac{E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}} = \frac{k/\sqrt{m}\ f_2(F_2^{-1}(p)) + O(1/m)}{1/\sqrt{m}\ \sqrt{var(\nu_m)}}$$

$$= \frac{kf_2(F_2^{-1}(p))}{\sqrt{var(\nu_m)}} + \frac{O(1/\sqrt{m})}{\sqrt{var(\nu_m)}}.$$

Now, $\sqrt{var(\nu_m)} < \infty\ \forall\ m$, and as shown earlier,

$$\lim_{m\to\infty} \sqrt{var(\nu_m)} = \sqrt{2(p)-2F(1,F_2^{-1}(p))}.$$

Therefore,

$$\frac{\mu^{(k)}}{\sigma(k)} \equiv \lim_{m\to\infty} \frac{E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}} = k\,\frac{f_2(F_2^{-1}(p))}{\sqrt{2p-2F(1,F_2^{-1}(p))}}.$$

Notice that this last result implies that the sequence $\{E\{T_m^{(k)}\}/\sigma_{T_m}^{(k)}\}$ $m = 1,2,\ldots,$ is bounded for all but perhaps a finite number of values of m.

Let $\Phi_m(z)$ = the cdf of $\dfrac{T_m^{(k)} - E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}}$. From before we have that $\Phi_m(z) \xrightarrow[m\to\infty]{} \Phi(z)$ $\forall\, z \in \mathbb{R}^1$, where Φ is the cdf of N(0,1). Since Φ is the cdf of a continuous distribution and $\{-E\{T_m^{(k)}\}/\sigma_{T_m}^{(k)}\}$ is a bounded sequence converging to the finite value $-\mu(k)/\sigma(k)$, we have (see Randles and Wolfe (1979), Theorem A.3.6)

$$\lim_{m\to\infty} \Phi_m\left(\frac{-E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}}\right) = \Phi\left(\frac{-\mu(k)}{\sigma(k)}\right) .$$

Therefore,

$$\lim_{m\to\infty} P(U_m^{(k)} < k) = \lim_{m\to\infty} P(T_m^{(k)} > 0)$$

$$= \lim 1 - \Phi_m\left(\frac{-E\{T_m^{(k)}\}}{\sigma_{T_m}^{(k)}}\right) = 1 - \Phi\left(\frac{-\mu(k)}{\sigma(k)}\right) = \Phi\left(\frac{\mu(k)}{\sigma(k)}\right),$$

by the symmetry of N(0,1) about zero. Thus, we have established that

$$\lim_{m\to\infty} P(U_m^{(k)} < k) = \Phi\left(k\, \frac{f_2(F_2^{-1}(p))}{\sqrt{2p - 2F(1, F_2^{-1}(p))}}\right)$$

which implies that

$$\lim_{m\to\infty} P\left(U_m^{(k)} \cdot \frac{f_2(F_2^{-1}(p))}{\sqrt{2p-2F(1,F_2^{-1}(p))}} < k\right) = \Phi(k),$$

and since this is true for all k, we conclude that

$$U_m^{(k)} \frac{f_2(F_2^{-1}(p))}{\sqrt{2p-2F(1,F_2^{-1}(p))}} \xrightarrow[m\to\infty]{L} N(0,1).$$

Recalling that $U_m^{(k)} = \sqrt{m}[\hat{F}_2^{-1}(\hat{p})-F_2^{-1}(p)]$ we subsequently conclude that

$$\sqrt{m}(\hat{F}_2^{-1}(\hat{p})-F_2^{-1}(p)) \xrightarrow[m\to\infty]{L} N\left(0, \frac{2p-2F(1,F_2^{-1}(p))}{[f_2(F_2^{-1}(p))]^2}\right).$$

This complete the proof of the lemma.

Theorem 1: Let (X_{11},X_{21}), (X_{12},X_{22}), ..., (X_{1m},X_{2m}) and (Y_{11},Y_{21}), (Y_{12},Y_{22}), ..., (Y_{1n},Y_{2n}) be independent random samples from the bivariate distributions dF and dG, respectively. Let X_1 and Y_1 be Bernoulli random variables with $Pr\{X_{1i}=1\} = p$, $i = 1(1)m$ and $Pr\{Y_{1j}=1\} = q$, $j = 1(1)n$. Let the marginal distributions of X_2 and Y_2 be denoted by $F_2(x_2)$ and $G_2(y_2)$, respectively. Assume that F_2 and G_2 and F_2^{-1} are continuous, possessing finite derivatives of up to at least third order. Let $F(1,\xi) \equiv \int_{\{X_1=1,X_2\leq\xi\}} dF$ and $G(1,\xi) \equiv \int_{\{Y_1=1,Y_2\leq\xi\}} dG$.

Assume that $F(1,\xi)$ and $G(1,\xi)$ are both continuous functions of ξ and also assume that $F(1,F_2^{-1}(p)) \neq p$.

Define

$$r \equiv \sum_{i=1}^{m} X_{1i}, \quad \hat{p} = r/m, \quad \hat{q} = \sum_{j=1}^{n} Y_{1j}/n$$

and define $\hat{F}_2^{-1}(\hat{p}) = X_2(r;m)$ = the r-th ordered value of $\{X_{21},X_{22},\ldots,X_{2m}\}$, where we shall define $X_2(0;m) = -\infty$ and $X_2(m+1,m) = \infty$. Finally, define

$$\hat{G}_2(k) \equiv \sum_{j=1}^{n} I(k-y_{2j})/n,$$

where $I(z) = 1$ if $z \geq 0$ and $I(z) = 0$ otherwise.

Suppose that n and m proceed to infinity at the same rate. Then, the statistic

$$\sqrt{n}[\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q} - (G_2F_2^{-1}(p)-q)]$$

is asymptotically normal with mean zero and finite variance.

Proof: Consider the decomposition

$$\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q} - (G_2F_2^{-1}(p)-q)$$

$$= [\hat{G}_2F_2^{-1}(p) - \hat{q} - (G_2F_2^{-1}(p)-q)] + [G_2\hat{F}_2^{-1}(\hat{p}) - G_2F_2^{-1}(p)]$$

$$+ [(\hat{G}_2\hat{F}_2^{-1}(\hat{p})-G_2\hat{F}_2^{-1}(\hat{p})) - (\hat{G}_2F_2^{-1}(p)-G_2F_2^{-1}(p))]$$

$$= I + II + III, \text{ say.}$$

Note that I is a function solely of $\{(Y_{1j},Y_{2j})\}$, $j = 1(1)n$, and II is a function solely of $\{(X_{1i},X_{2i})\}$, $i = 1(1)m$, so that I and II are independent. Therefore, to prove the theorem it suffices to show that

a) $\sqrt{n}I$ is asymptotically normal with mean zero and finite variance,

b) $\sqrt{m}II$ is asymptotically normal with mean zero and finite variance,

c) $\sqrt{n}III$ converges to zero in probability.

(See, for example, Theorems A.3.13 and A.3.15 from Randles and Wolfe (1979).)

(a) $\sqrt{n}I$ is Asymptotically Normal

$$\hat{G}_2F_2^{-1}(p) - \hat{q} = \sum_{j=1}^{n} I(F_2^{-1}(p)-Y_{2j})/n - \sum_{j=1}^{n} Y_{1j}/n$$

$$= \sum_{j=1}^{n} [I(F_2^{-1}(p)-Y_{2j}) - Y_{1j}]/n$$

$$= , \text{ say, } \sum_{j=1}^{n} \nu_j/n.$$

Observe that since the bivariate random variables $\{Y_{1j},Y_{2j}\}$ $j = 1(1)N$, are iid, it follows that the ν_j, $j = 1,2,\ldots,n$, are iid. Also;

$$E(\nu_j) = E[I(F_2^{-1}(p)-Y_{2j}) - Y_{1j}]$$

$$= G_2(F_2^{-1}(p)) - q,$$

$$\text{cov}\{I(F_2^{-1}(p)-Y_{2j}),\ Y_{1j}\}$$

$$= E\{I(F_2^{-1}(p)-Y_{2j})Y_{1j}\} - E\{I(F_2^{-1}(p)-Y_{2j})\}E\{Y_{1j}\}$$

$$= G(1,F_2^{-1}(p)) - qG_2(F_2^{-1}(p)),$$

so that,

$$\text{var}(\nu_j) = \text{var}[I(F_2^{-1}(p)-Y_{2j})] + \text{var}[Y_{1j}]$$

$$- 2\text{cov}\{I(F_2^{-1}(p)-Y_{2j}),\ Y_{1j}\}$$

$$= G_2(F_2^{-1}(p))[1-G_2(F_2^{-1}(p))] + q[1-q]$$

$$- 2\{G(1,F_2^{-1}(p)) - qG_2(F_2^{-1}(p))\}.$$

Obviously, $\text{var}(\nu_j) < \infty$ $\forall$ j and $\forall$ n, and so we have by the Lindeberg-Levy Central Limit Theorem that

$$\sqrt{n}\left(\sum_{j=1}^{n} \nu_j/n - E(\nu_j)\right) \xrightarrow[n\to\infty]{L} N(0,\text{var}(\nu_j))$$

that is,

$$\sqrt{n}[\hat{G}\hat{F}_2^{-1}(p) - \hat{q} - (G_2(F_2^{-1}(p)) - q] \xrightarrow[n\to\infty]{L} N(0,\mathrm{var}(v_j)).$$

This proves part (a).

(b) $\sqrt{m}II$ is Asymptotically Normal

$$\sqrt{m}II = \sqrt{m}[G_2\hat{F}_2^{-1}(\hat{p}) - G_2F_2^{-1}(p)].$$

Since G_2 is assumed to be a continuous function, asymptotic normality follows as a consequence of Lemma 1 by way of the "Delta-Method" (see, for example, Bishop et al. (1975)).

By Lemma 1,

$$\sqrt{m}[\hat{F}_2^{-1}(\hat{p})-F_2^{-1}(p)] \xrightarrow[m\to\infty]{L} N\left[0, \frac{2p-2F(1,F_2^{-1}(p))}{[f_2(F_2^{-1}(p))]^2}\right].$$

By the Delta-Method it then follows that

$$\sqrt{m}[G_2\hat{F}_2^{-1}(\hat{p})-G_2F_2^{-1}(p)] \xrightarrow[m\to\infty]{L} N\left[0, \frac{2p-2F(1,F_2^{-1}(p))}{[f_2(F_2^{-1}(p))]^2}\,[G_2'(F_2^{-1}(p))]^2\right].$$

Letting $g_2 = G_2'$, we conclude that

$$\sqrt{m}II \xrightarrow[m\to\infty]{L} N\left[0, \left[\frac{g_2(F_2^{-1}(p))}{f_2(F_2^{-1}(p))}\right]^2 \cdot [2p-2F(1,F_2^{-1}(p))]\right].$$

(c) $\underline{\sqrt{n}III = o_p(1)}$

$$\sqrt{n}III = \sqrt{n}[(\hat{G}_2\hat{F}_2^{-1}(\hat{p})-G_2\hat{F}_2^{-1}(\hat{p})) - (\hat{G}_2F_2^{-1}(p)-G_2F_2^{-1}(p))]$$

=, say, $\sqrt{n}[\omega_n]$. Note that $E[(\hat{G}_2\hat{F}_2^{-1}(\hat{p})-G_2\hat{F}_2^{-1}(\hat{p})) - (\hat{G}_2F_2^{-1}(p)-G_2F_2^{-1}(p))]$ $= EE\{\hat{G}_2\hat{F}_2^{-1}(\hat{p})-G_2\hat{F}_2^{-1}(\hat{p})\,|\,\hat{F}_2^{-1}(\hat{p})\} - E(\hat{G}_2F_2^{-1}(p)-G_2F_2^{-1}(p)) = 0$, by the unbiasedness of sample proportions. Thus, by Chebyshev's Inequality we have, $\forall\ \varepsilon > 0$,

$$P(|\sqrt{n}\omega_n| > \varepsilon) < \frac{n\mathrm{var}(\omega_n)}{\varepsilon^2},$$

so that to show $\sqrt{n}\omega_n = o_p(1)$ it is enough to show that $\mathrm{var}(\omega_n) = O(1/n^2)$ for then $n\mathrm{var} = O(1/n)$.

By a well-known result,

$$\mathrm{var}(\omega_n) = \mathrm{var}\{E(\omega_n|\hat{F}_2^{-1}(\hat{p}))\} + E\{\mathrm{var}(\omega_n|\hat{F}_2^{-1}(\hat{p}))\}.$$

Now,

$$E(\omega_n|\hat{F}_2^{-1}(\hat{p}))$$

$$= E(\hat{G}_2\hat{F}_2^{-1}(\hat{p})-G_2\hat{F}_2^{-1}(\hat{p})\,|\,\hat{F}_2^{-1}(\hat{p}))$$

$$- E(\hat{G}_2F_2^{-1}(p)-G_2F_2^{-1}(p)\,|\,\hat{F}_2^{-1}(\hat{p})) = 0$$

since for each fixed value of $\hat{F}_2^{-1}(\hat{p})$ the two expectations are zero owing, again, to the unbiasedness of sample proportions. Therefore,

$$\text{varE}(\omega_n|\hat{F}_2^{-1}(\hat{p})) = 0.$$

$$\text{Var}(\omega_n|\hat{F}_2^{-1}(\hat{p}))$$

$$= \text{var}(\hat{G}_2\hat{F}_2^{-1}(\hat{p})-G_2\hat{F}_2^{-1}(\hat{p})-\hat{G}_2F_2^{-1}(p)+G_2F_2^{-1}(p)|\hat{F}_2^{-1}(\hat{p}))$$

$$= \text{var}(\hat{G}_2\hat{F}_2^{-1}(\hat{p})-\hat{G}_2F_2^{-1}(p)|\hat{F}_2^{-1}(\hat{p}))$$

$$= \text{, say, } \text{var}(\hat{G}_2(\hat{\theta})-\hat{G}_2(\theta)|\hat{\theta})$$

where $\theta = F_2^{-1}(p)$, $\hat{\theta} = \hat{F}_2^{-1}(\hat{p})$. (Note that, with $\hat{F}_2^{-1}(\hat{p})$ fixed, $G_2(\hat{\theta})$ and $G_2(\theta)$ are constant and so have conditional variance equal to zero.) Thus,

$$E\{\text{var}(\omega_n|\hat{\theta})\} = E\{\frac{G_2(\hat{\theta})(1-G_2(\hat{\theta}))}{n} + \frac{G_2(\theta)(1-G_2(\theta))}{n}\}$$

$$- 2\ \text{cov}\{\hat{G}_2(\hat{\theta}),\hat{G}_2(\theta)\}.$$

Suppose now that $\hat{\theta} \geq \theta$, then

$$E\{\hat{G}_2(\hat{\theta})\hat{G}_2(\theta)\} = \frac{1}{n^2}\sum_{j=1}^{n} E\{I(\theta-Y_{2j})\} + \frac{1}{n^2}\sum\sum_{j\neq k} E\{I(\theta-Y_{2j})I(\hat{\theta}-Y_{2k})\}$$

$$= \frac{1}{n}G_2(\theta) + \frac{n-1}{n}G_2(\theta)G_2(\hat{\theta}).$$

Therefore,

$$\operatorname{cov}\{\hat{G}_2(\hat{\theta}),\hat{G}_2(\theta)\} = E\{\hat{G}_2(\hat{\theta})\hat{G}_2(\theta)\} - E\{\hat{G}_2(\theta)\}E\{\hat{G}_2(\theta)\}$$

$$= (G_2(\theta)(1-G_2(\hat{\theta})))/n$$

and the expectation becomes

$$E\{\frac{G_2(\hat{\theta})(1-G_2(\hat{\theta})}{n}\} + \frac{G_2(\theta)(1-G_2(\theta))}{n} - 2E\{\frac{G_2(\theta)(1-G_2(\hat{\theta}))}{n}\}$$

$$= \frac{1}{n}(E\{G_2(\hat{\theta})\} - E\{G_2(\hat{\theta})^2\} + G_2(\theta) - G_2(\theta)^2 - 2G_2(\theta) + 2G_2(\theta)E\{G_2(\hat{\theta})\})$$

$$= \frac{1}{n}(E\{G_2(\hat{\theta})\} - E\{G_2(\hat{\theta})^2\} - G_2(\theta) - G_2(\theta)^2 + 2G_2(\theta)E\{G_2(\hat{\theta})\}).$$

(2.2.4)

Expanding the functions $G_2(\cdot)$ and $G_2(\cdot)^2$ about θ gives (2.2.4)

$$= \frac{1}{n}[E(G_2(\theta) + G_2'(\theta)(\hat{\theta}-\theta) + G_2''(\theta)\frac{(\hat{\theta}-\theta)^2}{2} + o[(\hat{\theta}-\theta)^2])$$

$$- E(G_2(\theta)^2 + [G_2(\theta)^2]'(\hat{\theta}-\theta) + [G_2(\theta)^2]''\frac{(\hat{\theta}-\theta)^2}{2} + o[(\hat{\theta}-\theta)^2])$$

$$- G_2(\theta) - G_2(\theta)^2$$

$$+ 2G_2(\theta)E(G_2(\theta) + G_2'(\theta)(\hat{\theta}-\theta) + G_2''(\theta)\frac{(\hat{\theta}-\theta)^2}{2} + o[(\hat{\theta}-\theta)^2])].$$

Eventually we will show that $E(\hat{\theta}-\theta)^2 = O(1/n)$ and so we may ignore the $o(\hat{\theta}-\theta)^2$ terms for simplicity. Therefore we have, when $\hat{\theta} \geq \theta$,

$$E\{\mathrm{var}(\omega_n|\hat{\theta})\} = \frac{1}{n}(G_2'(\theta)E(\hat{\theta}-\theta) + G_2''(\theta)\frac{E(\hat{\theta}-\theta)^2}{2}$$

$$+ [G_2(\theta)^2]'E(\hat{\theta}-\theta) + [G_2(\theta)^2]''\frac{E(\hat{\theta}-\theta)^2}{2}$$

$$+ 2G_2(\theta)G_2'(\theta)E(\hat{\theta}-\theta) + 2G_2(\theta)G_2''(\theta)\frac{E(\hat{\theta}-\theta)^2}{2}$$

$$= \frac{1}{n}([G_2'(\theta) + [G_2(\theta)^2]' + 2G_2(\theta)G_2'(\theta)]E(\hat{\theta}-\theta)$$

$$+ \frac{[G_2''(\theta)+[G_2(\theta)^2]''+2G_2(\theta)G_2''(\theta)]E(\hat{\theta}-\theta)^2}{2}) .$$

On the other hand, if $\hat{\theta} < \theta$ then the covariance term is $(G_2(\hat{\theta})(1-G_2(\theta))/n$ and the expectation of the conditional variance is

$$\frac{1}{n}(EG_2(\hat{\theta}) - EG_2(\hat{\theta})^2 + G_2(\theta) - G_2(\theta)^2 - 2EG_2(\hat{\theta}) + 2G_2(\theta)EG_2(\hat{\theta}))$$

$$= \frac{1}{n}(-EG_2(\hat{\theta}) - EG_2(\hat{\theta})^2 + G_2(\theta) - G_2(\theta)^2 + 2G_2(\theta)EG_2(\hat{\theta}))$$

$$= \frac{1}{n}(-E[G_2(\theta) + G_2'(\theta)(\hat{\theta}-\theta) + G_2''(\theta)(\hat{\theta}-\theta)^2/2 + o[(\hat{\theta}-\theta)]^2]$$

$$- E[G_2(\theta)^2 + [G_2(\theta)^2]'(\hat{\theta}-\theta) + [G_2(\theta)^2]''(\hat{\theta}-\theta)^2/2 + o[(\hat{\theta}-\theta)^2]]$$

$$+ G_2(\theta) - G_2(\theta)^2 + 2G_2(\theta)E[G_2(\theta) + G_2'(\theta)(\hat{\theta}-\theta) + G_2''(\theta)(\hat{\theta}-\theta)^2/2$$

$$+ o[(\hat{\theta}-\theta)^2]]).$$

And so, ignoring the $o[(\hat{\theta}-\theta)^2]$ term, we have that when $\hat{\theta} < \theta$

$$E\{var(\omega_n|\hat{\theta})\}$$

$$= \frac{1}{n}(G_2'(\theta)E(\hat{\theta}-\theta) + G_2''(\theta)E(\hat{\theta}-\theta)^2/2 - [G_2(\theta)^2]'E(\hat{\theta}-\theta)$$

$$- [G_2(\theta)^2]''E(\hat{\theta}-\theta)^2/2 + 2G_2'(\theta)E(\hat{\theta}-\theta) + G_2''(\theta)E(\hat{\theta}-\theta)^2/2)$$

$$= \frac{1}{n}([G_2'(\theta) - [G_2(\theta)^2]' + 2G_2'(\theta)]E(\hat{\theta}-\theta)$$

$$+ 1/2[G_2''(\theta) - [G_2(\theta)^2]'' + G_2''(\theta)]E(\hat{\theta}-\theta)^2).$$

As all derivatives are assumed to be finite, to show that $var(\omega_n) = O(1/n^2)$ it will be enough to show that both $E(\hat{\theta}-\theta)$ and $E(\hat{\theta}-\theta)^2$ are of order 1/n.

Recall, $E(\hat{\theta}-\theta) = E(\hat{F}_2^{-1}(\hat{p}) - F_2^{-1}(p))$. We may reexpress the expectation as

$$= EE\{\hat{F}_2^{-1}(\hat{p}) - F_2^{-1}(p)|\hat{p}\}.$$

The conditional expectation may be recognized as a function of the expected value of the order statistic $\hat{F}_2^{-1}(\hat{p})$. Expanding this expected value (see David (1981) p. 80), we have

$$E\{\hat{F}_2^{-1}(\hat{p}) - F_2^{-1}(p)/\hat{p}\}$$

$$= F_2^{-1}(\hat{p}) + \frac{\hat{p}(1-\hat{p})}{2(n+2)}[F_2^{-1}(\hat{p})]'' + O(1/m^2) - F_2^{-1}(p).$$

Consequently, $E(\hat{\theta}-\theta)$

$$= E(F_2^{-1}(\hat{p}) - F_2^{-1}(p) + \frac{\hat{p}(1-\hat{p})}{2(n+2)} [\hat{F}_2^{-1}(\hat{p})]'' + O(1/m^2))$$

$$= E(F_2^{-1}(\hat{p}) - F_2^{-1}(p)) + O(1/m), \text{ since}$$

$(\hat{F}_2^{-1}(\hat{p}))''E\{\hat{p}(1-\hat{p})\}$ is bounded and $1/(2(n+2)) = O(1/m)$.

Now,

$$E(F_2^{-1}(\hat{p}) - F_2^{-1}(p))$$

$$= E(F_2^{-1}(p) + [F_2^{-1}(p)]'(\hat{p}-p) + \frac{[F_2^{-1}(p)]''}{2} (\hat{p}-p)^2 + o[(\hat{p}-p)^2] - F_2^{-1}(p))$$

$$= [F_2^{-1}(p)]'E(\hat{p}-p) + \frac{[F_2^{-1}(p)]''}{2} E(\hat{p}-p)^2 + o[(\hat{p}-p)^2].$$

But, $E(\hat{p}-p) = 0$ and $E(\hat{p}-p)^2 = p(1-p)/m = O(1/m)$. Therefore,

$$E(F_2^{-1}(\hat{p}) - F_2^{-1}(p)) = O(1/m)$$

$$\Rightarrow E(\hat{\theta}-\theta) = O(1/m) \quad (= O(1/n)).$$

Similarly,

$$E(\hat{\theta}-\theta)^2 = E(\hat{F}_2^{-1}(\hat{p}) - F_2^{-1}(p))^2$$

$$= EE\{(\hat{F}_2^{-1}(\hat{p}) - F_2^{-1}(p))^2/\hat{p}\}.$$

Considering $\hat{p}$ as fixed,

$$E(\hat{F}_2^{-1}(\hat{p}) - F_2^{-1}(p))^2$$

$$= E(F_2^{-1}(\hat{p}) - F_2^{-1}(p) + \frac{\hat{p}(1-\hat{p})}{2(n+2)} [\hat{F}_2^{-1}(\hat{p})]'' + O(1/m^2))^2$$

$$= E[F_2^{-1}(p) + [F_2^{-1}(p)]'(\hat{p}-p) + [F_2^{-1}(p)]'' \frac{(\hat{p}-p)^2}{2}$$

$$+ o[(\hat{p}-p)]^2 - F_2^{-1}(p) + \frac{\hat{p}(1-\hat{p})}{2(n+1)} [\hat{F}_2^{-1}(\hat{p})]'' + O(1/m^2)]^2.$$

Note that as each of the above terms are finite, after taking the square any term multipled by either of the terms

$$o[(\hat{p}-p)^2], \frac{[\hat{F}_2^{-1}(\hat{p})]''}{2(n+2)} p(1-p), \text{ or } O(1/m^2)$$

will be $O(1/m)$. Therefore, we may ignore these terms and noting that the two $F_2^{-1}(p)$ terms cancel we may write

$$E(\hat{\theta}-\theta)^2 = E([F_2^{-1}(p)]'(\hat{p}-p) + [F_2^{-1}(p)]''(\hat{p}-p)^2/2)^2$$

$$= [F_2^{-1}(p)]^2 E(\hat{p}-p)^2 + 2[F_2^{-1}(p)]' \frac{[F_2^{-1}(p)]''}{2} E(\hat{p}-p)^3$$

$$+ \frac{[F_2^{-1}(p)]''^2}{4} E(\hat{p}-p)^4.$$

For all $\hat{p}$, $0 < |\hat{p}-p| < 1$, and so it follows that $|(\hat{p}-p)^4| \leq |(\hat{p}-p)^3| \leq |(\hat{p}-p)^2|$. Therefore, $E(\hat{p}-p)^3 = O(E(\hat{p}-p)^2)$ and $E(\hat{p}-p)^4 = O(E(\hat{p}-p)^2)$. But, as is well-known $E(\hat{p}-p)^2 = O(1/m)$, and thus $E(\hat{\theta}-\theta)^2 = O(1/m)$.

Therefore, regardless of whether $\hat{\theta} < \theta$ or $\hat{\theta} \geq \theta$, it is the case that

$$E\{\text{var}(\omega_n|\hat{\theta})\} = \frac{1}{n} O(1/n) = O(1/n^2).$$

This implies that

$$\text{var}(\omega_n) = O(1/n^2), \text{ and thus}$$

$$\sqrt{n}III = o_p(1)$$

as was to be shown. This completes the proof of Theorem 1.

2.3 Comparing a Test Based Upon an S-nary Separator Variable with Several Tests Based Upon Continuous Separator Variables

We now consider extension of the statistic to the case when $T \geq 2$ and $S \geq 2$. Lemma 2 extends Lemma 1 to the multivariate distribution of S order statistics based upon orders r_s, $s = 1(1)S$, each of which has binomial distribution $BIN(m,p_s)$ where

$$p_s = \Pr\{X_1 \leq v_s\},$$

$v_1 < v_2 < \ldots < v_s$ representing the S values attainable by the discrete separator variable X_1.

Theorem 2 gives conditions for, and proves asymptotic multivariate normality of, the vector of differences in sensitivities between each test based upon a continuous separator variable and the test based upon a discrete separator after the test based upon the continuous separator has been set to have one of the specificities available to the test based upon the discrete separator. These differences are considered for each of the (T-1) tests based upon continuous separators and for each of the S specificities available to the test based upon the discrete separator.

Lemma 2: Let $\{\underline{X}_i\}$ i = 1(1)m be a random sample of m T-dimensional vectors where $\underline{X}_i' = (X_{1i}, X_{2i}, \ldots, X_{Ti})$ is such that X_{1i} is a random variable possessing $S < \infty$ possible values $v_1 < v_2 < \ldots < v_s$ with $\Pr\{X_1 \leq v_s\} = p_s$, s = 1(1)S. Also assume that for each t = 2,3,...,T X_t has marginal cdf F_t which is a continuous function of X_t possessing finite derivatives of up to at least third order with f_t denoting the first derivative.

Let $F_{1,t}$ denote the joint cdf of $\{X_1, X_t\}$ and assume that for each s, $F_{1,t}(v_s, \xi)$ is a continuous function of ξ. Define

$$r_s = \sum_{i=1}^{m} I(v_s - X_{1i}),$$

$$\hat{p}_s = r_s/m$$

and

$\hat{F}_t^{-1}(\hat{p}_s) = X_t(r_s;m)$ = the r_s-th ordered value of $\{X_{t1}, X_{t2}, \ldots, X_{tm}\}$, $t = 2,3,\ldots,T$, where we shall define $X_t(0;m) = -\infty$ and $X_t(m+1;m) = \infty$. Then, the distribution of

$$\sqrt{m}\,\mathrm{vec}\begin{bmatrix} \hat{F}_2^{-1}(\hat{p}_1)-F_2^{-1}(p_1) & \hat{F}_3^{-1}(\hat{p}_1)-F_3^{-1}(p_1) & \cdots & \hat{F}_T^{-1}(\hat{p}_1)-F_T^{-1}(p_1) \\ \hat{F}_2^{-1}(\hat{p}_2)-F_2^{-1}(p_2) & \hat{F}_3^{-1}(\hat{p}_2)-F_3^{-1}(p_2) & \cdots & \hat{F}_T^{-1}(\hat{p}_2)-F_T^{-1}(p_2) \\ \vdots & \vdots & & \vdots \\ \hat{F}_2^{-1}(\hat{p}_S)-F_2^{-1}(p_S) & \hat{F}_3^{-1}(\hat{p}_S)-F_3^{-1}(p_S) & \cdots & \hat{F}_T^{-1}(\hat{p}_S)-F_T^{-1}(p_S) \end{bmatrix}$$

converges in law to a (S(T-1)x1) multivariate normal distribution having mean $\underline{0}$, the (S(T-1)x1) vector of zeros, and (S(T-1)xS(T-1)) variance-covariance matrix

$$\underline{W}^{-1}\underline{\Omega}^{-1}\underline{\Sigma}^{X}\underline{\Omega}^{-1}\underline{W}^{-1}$$

where

$$\underline{W} = \text{diag}\left\{\begin{pmatrix} f_2(F_2^{-1}(p_1))/[2p_1-2F_{1,2}(v_1,F_2^{-1}(p_1))]^{1/2} \\ f_2(F_2^{-1}(p_2))/[2p_2-2F_{1,2}(v_2,F_2^{-1}(p_2))]^{1/2} \\ \vdots \\ f_2(F_2^{-1}(p_S))/[2p_S-2F_{1,2}(v_S,F_2^{-1}(p_S))]^{1/2} \\ \vdots \\ f_T(F_T^{-1}(p_1))/[2p_1-2F_{1,T}(v_1,F_T^{-1}(p_1))]^{1/2} \\ f_T(F_T^{-1}(p_2))/[2p_2-2F_{1,T}(v_2,F_T^{-1}(p_2))]^{1/2} \\ \vdots \\ f_T(F_T^{-1}(p_S))/[2p_S-2F_{1,T}(v_S,F_T^{-1}(p_S))]^{1/2} \end{pmatrix}\right\}$$

$$\underline{\Omega} = (\text{diag}\{\Sigma^X\})^{1/2}$$

and

$$\underline{\Sigma}^X = \begin{pmatrix} \sigma_{22}^{11} & \sigma_{22}^{12} & \cdots & \sigma_{22}^{1S} & \sigma_{23}^{11} & \sigma_{23}^{12} & \cdots & \sigma_{23}^{1S} & \cdots & \sigma_{2T}^{11} & \sigma_{2T}^{12} & \cdots & \sigma_{2T}^{1S} \\ \sigma_{22}^{21} & \sigma_{22}^{22} & \cdots & \sigma_{22}^{2S} & \sigma_{23}^{21} & \sigma_{23}^{22} & \cdots & \sigma_{23}^{2S} & \cdots & \sigma_{2T}^{21} & \sigma_{2T}^{22} & \cdots & \sigma_{2T}^{2S} \\ & & \ddots & & & & & & & & & & \\ \sigma_{22}^{S1} & \sigma_{22}^{S2} & \cdots & \sigma_{22}^{SS} & \sigma_{23}^{S1} & \sigma_{23}^{S2} & \cdots & \sigma_{23}^{SS} & \cdots & \sigma_{2T}^{S1} & \sigma_{2T}^{S2} & \cdots & \sigma_{2T}^{SS} \\ & & & & \ddots & & & & & & & & \\ & & & & & \ddots & & & & & & & \\ & & & & & & \ddots & & & & & & \\ \sigma_{T2}^{11} & \sigma_{T2}^{12} & \cdots & \sigma_{T2}^{1S} & \sigma_{T3}^{11} & \sigma_{T3}^{12} & \cdots & \sigma_{T3}^{1S} & \cdots & \sigma_{TT}^{11} & \sigma_{TT}^{12} & \cdots & \sigma_{TT}^{1S} \\ \sigma_{T2}^{21} & \sigma_{T2}^{22} & \cdots & \sigma_{T2}^{2S} & \sigma_{T3}^{21} & \sigma_{T3}^{22} & \cdots & \sigma_{T3}^{2S} & \cdots & \sigma_{TT}^{21} & \sigma_{TT}^{22} & \cdots & \sigma_{TT}^{2S} \\ & & & & & & & & \ddots & & & & \\ \sigma_{T2}^{S1} & \sigma_{T2}^{S2} & \cdots & \sigma_{T2}^{SS} & \sigma_{T3}^{S1} & \sigma_{T3}^{S2} & \cdots & \sigma_{T3}^{SS} & \cdots & \sigma_{TT}^{S1} & \sigma_{TT}^{S2} & \cdots & \sigma_{TT}^{SS} \end{pmatrix}$$

with diag $\{\underline{a}\}$ of an Nx1 vector being the NxN diagonal matrix having the elements of the vector $\underline{a}$ on the diagonal, diag $\{\underline{W}\}$ of an NxN matrix being the NxN diagonal matrix retaining the diagonal elements of $\underline{W}$, and where we define when $s_1 = s_2$ and $t_1 = t_2$

$$\sigma_{t_1 t_2}^{s_1 s_2} = 2p_{s_1} - 2F_{1t_1}(v_{s_1}, F_{t_1}^{-1}(p_{s_1}))$$

when $s_1 = s_2$ and $t_1 \neq t_2$

$$\sigma_{t_1 t_2}^{s_1 s_2} = p_{s_1} - F_{1t_1}(v_{s_1}, F_{t_1}^{-1}(p_{s_1})) - F_{1t_2}(v_{s_1}, F_{t_2}^{-1}(p_{s_1}))$$

$$+ F_{t_1 t_2}(F_{t_1}^{-1}(p_{s_1}), F_{t_2}^{-1}(p_{s_1}))$$

when $s_1 < s_2$ and $t_1 = t_2$

$$\sigma_{t_1 t_2}^{s_1 s_2} = 2p_{s_1} - F_{1t_1}(v_{s_1}, F_{t_1}^{-1}(p_{s_2})) - F_{1t_1}(v_{s_2}, F_{t_1}^{-1}(p_{s_1}))$$

and when $s_1 \neq s_2$ and $t_1 \neq t_2$

$$\sigma_{t_1 t_2}^{s_1 s_2} = p_{s*} - F_{1t_1}(v_{s_2}, F_{t_1}^{-1}(p_{s_1})) - F_{1t_2}(v_{s_1}, F_{t_2}^{-1}(p_{s_2}))$$

$$+ F_{t_1 t_2}(F_{t_1}^{-1}(p_{s_1}), F_{t_2}^{-1}(p_{s_2}))$$

where $s^* = \min(s_1, s_2)$.

Proof: Without loss of generality we prove the case when S = 2 and T = 2.

Let k_{21}, k_{22}, k_{31}, and k_{32} be real-valued constants. By an argument similar to that used in Lemma 1, it follows that

$$\Pr\left\{\sqrt{m}\begin{bmatrix}\hat{F}_2^{-1}(\hat{p}_1)-F_2^{-1}(p_1)\\ \hat{F}_2^{-1}(\hat{p}_2)-F_2^{-1}(p_2)\\ \hat{F}_3^{-1}(\hat{p}_1)-F_3^{-1}(p_1)\\ \hat{F}_3^{-1}(\hat{p}_2)-F_3^{-1}(p_2)\end{bmatrix} \leq \begin{bmatrix}k_{21}\\ k_{22}\\ k_{31}\\ k_{32}\end{bmatrix}\right\}$$

$$= \Pr\left\{\begin{bmatrix}\hat{F}_2(b_{21})-\hat{p}_1\\ \hat{F}_2(b_{22})-\hat{p}_2\\ \hat{F}_3(b_{31})-\hat{p}_1\\ \hat{F}_3(b_{32})-\hat{p}_2\end{bmatrix} \geq \begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix}\right\} \qquad (2.3.1)$$

where

$$b_{ts} = (k_{ts}/\sqrt{m} + F_t^{-1}(p_s)).$$

For m fixed and for i = 1(1)m, let

$$\underline{U}_{m_i} = \begin{bmatrix} I(b_{21}-X_{2i}) - I(v_1-X_{1i}) \\ I(b_{22}-X_{2i}) - I(v_2-X_{1i}) \\ I(b_{31}-X_{3i}) - I(v_1-X_{1i}) \\ I(b_{32}-X_{3i}) - I(v_2-X_{1i}) \end{bmatrix} .$$

By identicality when m is fixed,

$$\underline{\mu}_m = E(\underline{U}_{m_i}) = \begin{bmatrix} (\mu_m)_{21} \\ (\mu_m)_{22} \\ (\mu_m)_{31} \\ (\mu_m)_{32} \end{bmatrix} = \begin{bmatrix} F_2(b_{21})-p_1 \\ F_2(b_{22})-p_2 \\ F_3(b_{31})-p_1 \\ F_3(b_{32})-p_2 \end{bmatrix}$$

and

$$\text{var}(\underline{U}_{m_i}) = \underline{\Sigma}_m^X = \begin{bmatrix} (\sigma_m)_{22}^{11} & (\sigma_m)_{22}^{12} & (\sigma_m)_{23}^{11} & (\sigma_m)_{23}^{12} \\ & (\sigma_m)_{22}^{22} & (\sigma_m)_{23}^{21} & (\sigma_m)_{23}^{22} \\ & " & (\sigma_m)_{33}^{11} & (\sigma_m)_{33}^{12} \\ & & & (\sigma_m)_{33}^{22} \end{bmatrix}$$

where

$$(\sigma_m)^{s_1 s_2}_{t_1 t_2} = \text{cov}\{[I(b_{t_1 s_1} - X_{t_1 i}) - I(v_{s_1} - X_{1i})],\ [I(b_{t_2 s_2} - X_{t_2}) - I(v_{s_2} - X_{1i})]\}$$

$$= \text{cov}\{I(b_{t_1 s_1} - X_{t_1 i}),\ I(b_{t_2 s_2} - X_{t_2 i})\}$$

$$- \text{cov}\{I(b_{t_1 s_1} - X_{t_1 i}),\ I(v_{t_2} - X_{1i})\}$$

$$- \text{cov}\{I(v_{s_1} - X_{1i}),\ I(b_{t_2 s_2} - X_{t_2 i})\}$$

$$+ \text{cov}\{I(v_{s_1} - X_{1i}),\ I(v_{s_2} - X_{1i})\}.$$

The case when $t_1 = t_2$ and $s_1 = s_2$ is not considered here as this covariance is the variance term determined in Lemma 1.

Recall that $b_{ts} = k_{ts}/\sqrt{m} + F_t^{-1}(p_s)$. Since $p_2 > p_1$ we have that $F_t^{-1}(p_2) > F_t^{-1}(p_1)$ and thus, for whatever values of k_{t_1} and k_{t_2} may have been selected $\exists\ m_t^* \ni b_{t_2} \geq b_{t_1}$ whenever $m > m_t^*$. This is to say that, for large enough m, $b_{21} \leq b_{22}$ and $b_{31} \leq b_{32}$. Therefore, when $t_1 = t_2$ and $s_1 \neq s_2$ the covariance term is

$$E\{I(b_{t_1 1} - X_{t_1 i})\} - E\{I(b_{t_1 1} - X_{t_1 i})\}E\{I(b_{t_1 2} - X_{t_1 i})\}$$

$$- [E\{I(b_{t_1 1} - X_{t_1 i})I(v_2 - X_{1i})\} - E\{I(b_{t_1 1} - X_{t_1 i})\}E\{I(v_2 - X_{1i})\}]$$

$$- [E\{I(b_{t_1 2} - X_{t_1 i})I(v_1 - X_{1i})\} - E\{I(b_{t_1 2} - X_{t_1 i})\}E\{I(v_1 - X_{1i})\}]$$

$$+ [E\{I(v_1-X_{1i})\} - E\{I(v_1-X_{1i})\}E\{I(v_2-X_{1i})\}]$$

$$= F_{t_1}(b_{t_1 1}) - F_{t_1}(b_{t_1 1})F_{t_1}(b_{t_1 2})$$

$$- [F_{1t_1}(v_2,b_{t_1 1}) - p_2 F_{t_1}(b_{t_1 1})]$$

$$- [F_{1t_1}(v_1,b_{t_1 2}) - p_1 F_{t_1}(b_{t_1 2})]$$

$$+ [p_1-p_1p_2].$$

When $t_1 \neq t_2$ and $s_1 = s_2$, the covariance term is

$$E\{I(b_{t_1 s_1}-X_{t_1 i})I(b_{t_2 s_1}-X_{t_2 i})\} - E\{I(b_{t_1 s_1}-X_{t_1 i})\}E\{I(b_{t_2 s_1}-X_{t_2 i})\}$$

$$- [E\{I(b_{t_1 s_1}-X_{t_1 i})I(v_{s_1}-X_{1i})\} - E\{I(b_{t_1 s_1}-X_{t_1 i})\}E\{I(v_{s_1}-X_{1i})\}]$$

$$- [E\{I(b_{t_2 s_1}-X_{t_2 i})I(v_{s_1}-X_{1i})\} - E\{I(b_{t_2 s_1}-X_{t_2 i})\}E\{I(v_{s_1}-X_{1i})\}]$$

$$+ [E\{I(v_s-X_{1i})\} - E\{I(v_s-X_{1i}\}^2]$$

$$= F_{23}(b_{2s},b_{3s}) - F_2(b_{2s})F_3(b_{3s})$$

$$- [F_{12}(v_s,b_{2s}) - p_s F_2(b_{2s})]$$

$$- [F_{13}(v_s,b_{3s}) - p_s F_3(b_{3s})]$$

$$+ [p_s-p_s^2].$$

Finally, when $t_1 \neq t_2$ and $s_1 \neq s_2$ the covariance term is either

$$E\{I(b_{21}-X_{2i})I(b_{32}-X_{3i})\} - E\{I(b_{21}-X_{2i})\}E\{I(b_{32}-X_{3i})\}$$

$$- [E\{I(b_{21}-X_{2i})I(v_2-X_{1i})\} - E\{I(b_{21}-X_{2i})\}E\{I(v_2-X_{1i})\}]$$

$$- [E\{I(b_{32}-X_{3i})I(v_1-X_{1i})\} - E\{I(b_{32}-X_{3i})\}E\{I(v_1-X_{1i})\}]$$

$$+ [E\{I(v_1-X_{1i})I(v_2-X_{1i})\} - E\{I(v_1-X_{1i})\}E\{I(v_2-X_{1i})\}$$

$$= F_{23}(b_{21},b_{32}) - F_2(b_{21})F_3(b_{32})$$

$$- [F_{12}(v_2,b_{21}) - p_2F_2(b_{21})]$$

$$- [F_{13}(v_1,b_{32}) - p_1F_3(b_{32})]$$

$$+ [p_1-p_1p_2]$$

or

$$= F_{23}(b_{22},b_{31}) - F_2(b_{22})F_3(b_{31})$$

$$- [F_{12}(v_1,b_{22}) - p_1F_2(b_{22})]$$

$$- [F_{13}(v_2,b_{31}) - p_2F_3(b_{31})]$$

$$+ [p_1-p_1p_2].$$

Now, let $\underline{\mu} = \lim_{m\to\infty} \underline{\mu}_m$

$= \lim_{m\to\infty}[F(b_{21})-p_1,\ F(b_{22})-p_2,\ F(b_{31})-p_1,\ F(b_{32})-p_2]'$

$= [0,0,0,0]'$, and let

$$\underline{\Sigma}^X = \lim_{m\to\infty} \Sigma_m^X = [\sigma_{t_1t_2}^{s_1s_2}]$$

where

$$\sigma_{t_1t_2}^{s_1s_2} = \lim_{m\to\infty} (\sigma_m)_{t_1t_2}^{s_1s_2} .$$

Recalling that $b_{ts} = k_{ts}/\sqrt{m} + F_t^{-1}(p_s)$ we have $\lim_{m\to\infty} b_{ts} = F_t^{-1}(p_s)$ so that, as in Lemma 1,

$$\sigma_{22}^{11} = 2p_1 - 2F_{12}(v_1,F_2^{-1}(p_1))$$

$$\sigma_{22}^{22} = 2p_2 - 2F_{12}(v_2,F_2^{-1}(p_2))$$

$$\sigma_{33}^{11} = 2p_1 - 2F_{13}(v_1,F_3^{-1}(p_1))$$

$$\sigma_{33}^{22} = 2p_2 - 2F_{13}(v_2,F_3^{-1}(p_2))$$

and now, in addition,

$$\sigma_{22}^{12} = p_1 - p_1p_2 - F_{12}(v_2, F_2^{-1}(p_1)) + p_1p_2 - F_{12}(v_1, F_2^{-1}(p_2))$$

$$+ p_1p_2 + p_1 - p_1p_2$$

$$= 2p_1 - F_{12}(v_1, F_2^{-1}(p_2)) - F_{12}(v_2, F_2^{-1}(p_1))$$

$$\sigma_{33}^{12} = p_1 - p_1p_2 - F_{13}(v_2, F_3^{-1}(p_1)) + p_1p_2 - F_{13}(v_1, F_3^{-1}(p_2))$$

$$+ p_1p_2 + p_1 - p_1p_2$$

$$= 2p_1 - F_{13}(v_1, F_3^{-1}(p_2)) - F_{13}(v_2, F_3^{-1}(p_1))$$

$$\sigma_{23}^{11} = F_{23}(F_2^{-1}(p_1), F_3^{-1}(p_1)) - p_1^2 - F_{12}(v_1, F_2^{-1}(p_1)) + p_1^2$$

$$- F_{13}(v_1, F_3^{-1}(p_1)) + p_1^2 + p_1 - p_1^2$$

$$= p_1 - F_{12}(v_1, F_2^{-1}(p_1)) - F_{13}(v_1, F_3^{-1}(p_1)) + F_{23}(F_2^{-1}(p_1), F_3^{-1}(p_1))$$

$$\sigma_{23}^{22} = F_{23}(F_2^{-1}(p_2), F_3^{-1}(p_2)) - p_2^2 - F_{12}(v_2, F_2^{-1}(p_2)) + p_2^2$$

$$- F_{13}(v_2, F_3^{-1}(p_2)) + p_2^2 + p_2 - p_2^2$$

$$= p_2 - F_{12}(v_2, F_2^{-1}(p_2)) - F_{13}(v_2, F_3^{-1}(p_2)) + F_{23}(F_2^{-1}(p_2), F_3^{-1}(p_2))$$

$$\sigma_{23}^{12} = F_{23}(F_2^{-1}(p_1), F_3^{-1}(p_2)) - p_1p_2 - F_{12}(v_2, F_2^{-1}(p_1))$$

$$+ p_1p_2 - F_{13}(v_1, F_3^{-1}(p_2)) + p_1p_2 + p_1 - p_1p_2$$

$$= p_1 - F_{12}(v_2, F_2^{-1}(p_1)) - F_{13}(v_1, F_3^{-1}(p_2)) + F_{23}^{-1}(F_2^{-1}(p_1), F_3^{-1}(p_2))$$

and

$$\sigma_{23}^{21} = F_{23}(F_2^{-1}(p_2), F_3^{-1}(p_1)) - p_1p_2 - F_{12}(v_1, F_2^{-1}(p_2))$$

$$+ p_1p_2 - F_{13}(v_2, F_3^{-1}(p_1)) + p_1p_2 + p_1 - p_1p_2$$

$$= p_1 - F_{12}(v_1, F_2^{-1}(p_2)) - F_{13}(v_2, F_3^{-1}(p_1)) + F_{23}(F_2^{-1}(p_2), F_3^{-1}(p_1)).$$

Let $\underline{T}_m = \frac{1}{m} \sum_{i=1}^{m} \underline{U}_{m_i}$. We note that the righthand side of equation (2.3.1) may be written as $\Pr\{\underline{T}_m \geq 0\}$. We now show that $\underline{T}_m$ is asymptotically normal.

Let $\underline{\lambda}' = [\lambda_1, \lambda_2, \lambda_3, \lambda_4]$ be a vector of real, nonzero constants.

$$\underline{\lambda}'\underline{T}_m = \frac{1}{m} \sum_{i=1}^{m} \underline{\lambda}'\underline{U}_{m_i} = \frac{1}{m} \sum_{i=1}^{n} \nu_{m_i}, \text{ say.}$$

Let $\bar{\mu}_m = E(\nu_{m_i})$. We have, for m fixed,

$$\bar{\mu}_m = \underline{\lambda}'\underline{\mu}_m \quad \text{and} \quad \text{var}(\nu_{m_i}) = \underline{\lambda}'\underline{\Sigma}^X\underline{\lambda}.$$

To show the asymptotic normality of $\underline{T}_m$ it suffices to show that

$$\sqrt{m}\left(\frac{1}{m} \sum_{i=1}^{m} \nu_{m_i} - \bar{\mu}_m\right) \xrightarrow{L} N(0, \underline{\lambda}'\underline{\Sigma}^X\underline{\lambda}).$$

As in the proof of Lemma 1, identicality of the ν_{m_i} when m is fixed gives the Lindeberg-Condition to be, given $\eta > 0$,

$$\lim_{m\to\infty} \frac{1}{\mathrm{var}(\nu_m)} \int_{\left\{\left|\frac{\nu-\bar{\mu}_m}{\sigma_m}\right| > \eta\right\}} (\nu-\bar{\mu}_m)^2 dF_{\nu_m} = 0,$$

where $\sigma_m = \sqrt{m\mathrm{var}(\nu_m)}$ and F_{ν_m} represents the common distribution of $\{\nu_{m_i}\}$, i = 1(1)m.

By the previous work, it follows that $\lim_{m\to\infty} \frac{1}{\mathrm{var}(\nu_m)}$ exists and is finite.

Now,

$$(\nu-\bar{\mu}_m)^2 = (\nu_{m_i}-\bar{\mu}_m)^2 \text{ for } i = 1(1)m$$

$$= (\lambda_1[I(b_{21}-X_{2i}) - I(v_1-X_{1i}) - (F_2(b_{21})-p_1)]$$

$$+ \lambda_2[I(b_{22}-X_{2i}) - I(v_2-X_{1i}) - (F(b_{22})-p_2)]$$

$$+ \lambda_3[I(b_{31}-X_{3i}) - I(v_1-X_{1i}) - (F(b_{31})-p_1)]$$

$$+ \lambda_4[I(b_{32}-X_{3i}) - I(v_2-X_{1i}) - (F(b_{32}-p_2)])^2.$$

As the maximum value of any of the quantities multiplying the λ's is 2, we have

$$(\nu-\bar{\mu}_m)^2 \leq 4\lambda_1^2 + 4\lambda_2^2 + 4\lambda_3^2 + 4\lambda_4^2 + 2\cdot|\lambda_1\lambda_2|\cdot 4$$

$$+ 2\cdot|\lambda_1\lambda_3|\cdot 4 + 2\cdot|\lambda_1\lambda_4|\cdot 4 + 2\cdot|\lambda_2\lambda_3|\cdot 4$$

$$+ 2\cdot|\lambda_3\lambda_4|\cdot 4 < \infty.$$

Therefore, to establish the Lindeberg-Condition it suffices to show that

$$\lim_{m\to\infty} \int_{\left\{\left|\frac{\nu-\mu_m}{\sigma_m}\right| > \eta\right\}} dF_{\nu_m} = 0.$$

But, by Chebyshev's Inequality,

$$P\left\{\left|\frac{\nu-\bar{\mu}_m}{\sigma_m}\right| > \eta\right\} \leq \frac{\mathrm{var}(\nu_m)}{\sigma_m^2\eta^2} = \frac{1}{m\eta^2} \longrightarrow 0 \quad \text{as} \quad m\to\infty .$$

Thus,

$$\sqrt{m}\left(\frac{1}{m}\sum_{i=1}^{m} \nu_{m_i} - \bar{\mu}_m\right) \xrightarrow{L} N(0,\sigma_\infty^2)$$

where

$$\sigma_\infty^2 = \lim_{m\to\infty} \sigma_m = \lim_{m\to\infty} \mathrm{var}(\underline{\lambda}'\underline{U}_m)$$

$$= \lim_{m\to\infty} \underline{\lambda}'\underline{\Sigma}^{X}_{m}\underline{\lambda} = \underline{\lambda}'\underline{\Sigma}^{X}\underline{\lambda} .$$

Since this is true for any $\underline{\lambda}$, it follows that

$$\underline{T}_m \xrightarrow{L} N(\underline{0},\Sigma^X).$$

As demonstrated in Lemma 1,

$$\lim_{m\to\infty} \frac{(\mu_m)_{ts}}{(\frac{1}{m}\sigma^{ss}_{tt})^{1/2}} = k_{ts} \frac{f_t(F_t^{-1}(p_s))}{[2p_s - 2F_{1t}(v_s, F_t^{-1}(p_s))]^{1/2}}.$$

Let

$$\underline{\gamma}'_m = \left[\frac{(\mu_m)_{21}}{(\frac{1}{m}(\sigma_m)^{11}_{22})^{1/2}}, \frac{(\mu_m)_{22}}{(\frac{1}{m}(\sigma_m)^{22}_{22})^{1/2}}, \frac{(\mu_m)_{31}}{(\frac{1}{m}(\sigma_m)^{11}_{33})^{1/2}}, \frac{(\mu_m)_{32}}{(\frac{1}{m}(\sigma_m)^{22}_{33})^{1/2}} \right]$$

and

$$\underline{\gamma}' = \lim_{m\to\infty} \underline{\gamma}'_m$$

and recalling equation (2.3.1) we have

$$\lim_{m\to\infty} \Pr \left\{ \sqrt{m} \begin{bmatrix} \hat{F}_2^{-1}(\hat{p}_1) - F_2^{-1}(p_1) \\ \hat{F}_2^{-1}(\hat{p}_2) - F_2^{-1}(p_2) \\ \hat{F}_3^{-1}(\hat{p}_1) - F_3^{-1}(p_1) \\ \hat{F}_3^{-1}(\hat{p}_2) - F_3^{-1}(p_2) \end{bmatrix} < \begin{bmatrix} k_{21} \\ k_{22} \\ k_{31} \\ k_{32} \end{bmatrix} \right\}$$

$$= \lim_{m\to\infty} \Pr\left\{(-1)\begin{bmatrix} \dfrac{(\hat{F}_2(b_{21})-\hat{p}_1) - (\mu_m)_{21}}{(\frac{1}{m}(\sigma_m)_{22}^{11})^{1/2}} \\ \dfrac{(\hat{F}_2(b_{22})-\hat{p}_2) - (\mu_m)_{22}}{(\frac{1}{m}(\sigma_m)_{22}^{22})^{1/2}} \\ \dfrac{(\hat{F}_3(b_{31})-\hat{p}_1)-(\mu_m)_{31}}{(\frac{1}{m}(\sigma_m)_{33}^{11})^{1/2}} \\ \dfrac{(\hat{F}_3(b_{32})-\hat{p}_2)-(\mu_m)_{32}}{(\frac{1}{m}(\sigma_m)_{33}^{22})^{1/2}} \end{bmatrix} \leq \underline{\gamma}_m\right\}. \qquad (2.3.2)$$

Let

$$\underline{\Omega}_m = \text{diag}\begin{bmatrix} ((\sigma_m)_{22}^{11})^{1/2} \\ ((\sigma_m)_{22}^{22})^{1/2} \\ ((\sigma_m)_{33}^{11})^{1/2} \\ ((\sigma_m)_{33}^{22})^{1/2} \end{bmatrix}$$

and

$$\underline{\Omega} = \lim_{m\to\infty} \underline{\Omega}_m$$

then,

$$(2.3.2) = \lim_{m\to\infty} P\{-\underline{\Omega}_m^{-1}(\underline{T}_m - \underline{\mu}_m) \leq \underline{\gamma}_m\}.$$

By the asymptotic normality of T_m

$$\lim_{m\to\infty} P\{-\underline{\Omega}_m^{-1}(\underline{T}_m - \underline{\mu}_m) \leq \underline{\gamma}_m\} = \underline{\Phi}(\underline{\gamma})$$

where $\underline{\Phi}(\ )$ represents the multivariate $N_4(\underline{0}, \underline{\Omega}^{-1}\underline{\Sigma}^X\underline{\Omega}^{-1})$ cdf. Let

$$\underline{W} = \text{diag}\left[\begin{array}{c} \dfrac{f_2(F_2^{-1}(p_1))}{[2p_1 - 2F_{12}(v_1, F_2^{-1}(p_1))]^{1/2}} \\ \\ \dfrac{f_2(F_2^{-1}(p_2))}{[2p_2 - 2F_{12}(v_2, F_2^{-1}(p_2))]^{1/2}} \\ \\ \dfrac{f_3(F_3^{-1}(p_1))}{[2p_1 - 2F_{13}(v_1, F_3^{-1}(p_1))]^{1/2}} \\ \\ \dfrac{f_3(F_3^{-1}(p_2))}{[2p_2 - 2F_{13}(v_2, F_3^{-1}(p_2))]^{1/2}} \end{array}\right].$$

By an argument similar to that found in Lemma 1 we may thus conclude that

$$\sqrt{m}\begin{bmatrix} \hat{F}_2^{-1}(\hat{p}_1) - F_2^{-1}(p_1) \\ \hat{F}_2^{-1}(\hat{p}_2) - F_2^{-1}(p_2) \\ \hat{F}_3^{-1}(\hat{p}_1) - F_3^{-1}(p_1) \\ \hat{F}_3^{-1}(\hat{p}_2) - F_3^{-1}(p_2) \end{bmatrix} \xrightarrow{L} N_4(\underline{0}, \underline{W}^{-1}\underline{\Omega}^{-1}\underline{\Sigma}^{X}\underline{\Omega}^{-1}\underline{W}^{-1})$$

thus proving the lemma.

Theorem 2: Let $\{\underline{X}_i\}$ i = 1(1)m and $\{\underline{Y}_j\}$ j = 1(1)n be independent random samples of T-dimensional real-valued vectors. Let F_t represent the marginal distribution of X_{ti} and G_t represent the marginal distribution of Y_j possessing the following properties:

(a) F_1 and G_1 are discrete cdfs having support confined to the finite set $\{v_1 < v_2 < \ldots < v_S\}$

and

(b) for each t = 2(1)T, F_t and G_t are continuous functions possessing finite derivatives of up to at least third order, with first derivatives denoted by f_t and g_t, respectively.

Let $F_{t_1t_2}$ represent the joint distribution of (X_{t_1}, X_{t_2}) and $G_{t_1t_2}$ represent the joint distribution of (Y_{t_1}, Y_{t_2}) where it is required that $\forall$ s = 1(1)S, $F_{1t_2}(v_s,\xi)$ and $G_{1t_2}(v_s,\xi)$ are continuous functions of ξ, t_2 = 2(1)T.

Also, let $p_s = F_1(v_s)$, $q_s = G_1(v_s)$ and define

$$r_s = \sum_{i=1}^{m} I(v_s - X_{1i})$$

$$\hat{p}_s = r_s/m$$

$\hat{F}_t^{-1}(\hat{p}_s) = X_t(r_s;m)$ = the r_s-th ordered value of $\{X_{t1}, X_{t2}, \ldots, X_{tm}\}$, $t = 2(1)T$ where we define $X_t(0;m) = -\infty$ and $X_t(m+1;m) = +\infty$ and

$$I(Z) = 1 \quad \text{if} \quad Z \geq 0$$
$$0 \quad \text{if} \quad Z < 0.$$

Finally, define

$$\hat{q}_s = \frac{1}{n} \sum_{j=1}^{n} I(v_s - Y_{1j})$$

and

$$\hat{G}_t(\xi) = \frac{1}{n} \sum_{j=1}^{n} I(\xi - Y_{tj}).$$

Then,

$$\sqrt{m}\,\mathrm{vec}\begin{bmatrix} [(\hat{G}_2\hat{F}_2^{-1}(\hat{p}_1)-\hat{q}_1)-(G_2F_2^{-1}(p_1)-q_1)] & \cdots & [(\hat{G}_T\hat{F}_T^{-1}(\hat{p}_1)-\hat{q}_1)-(G_TF_T^{-1}(p_1)-q_1)] \\ \vdots & & \vdots \\ [(\hat{G}_2\hat{F}_2^{-1}(\hat{p}_s)-\hat{q}_s)-(G_2F_2^{-1}(p_s)-q_s)] & \cdots & [(\hat{G}_T\hat{F}_T^{-1}(\hat{p}_s)-\hat{q}_s)-(G_TF_T^{-1}(p_s)-q_s)] \end{bmatrix}$$

has an asymptotic distribution that is multivariate normal with mean $\underline{0}$ (S(T-1)x1) and variance-covariance matrix

$$\underline{\Sigma}^{Y} + \underline{g}\,\underline{W}^{-1}\,\underline{\Omega}^{-1}\underline{\Sigma}^{X}\underline{\Omega}^{-1}\underline{W}^{-1}\underline{g}$$

where

$\underline{W}$, $\underline{\Omega}$ and $\underline{\Sigma}^{X}$ are as given in Lemma 2

$$\underline{g} = \operatorname{diag}\begin{pmatrix} g_2(F_2^{-1}(p_1)) \\ g_2(F_2^{-1}(p_2)) \\ \vdots \\ g_2(F_2^{-1}(p_S)) \\ \vdots \\ g_T(F_T^{-1}(p_1)) \\ g_T(F_T^{-1}(p_2)) \\ \vdots \\ g_T(F_T^{-1}(p_S)) \end{pmatrix}$$

and

$$\underline{\Sigma}^{Y} = \begin{pmatrix} \Gamma_{22}^{11}\Gamma_{22}^{12} & \cdots & \Gamma_{22}^{1S}\Gamma_{23}^{11}\Gamma_{23}^{12} & \cdots & \Gamma_{23}^{1S} & \cdots & \Gamma_{2T}^{11}\Gamma_{2T}^{12} & \cdots & \Gamma_{2T}^{1S} \\ \Gamma_{22}^{21}\Gamma_{22}^{22} & \cdots & \Gamma_{22}^{2S}\Gamma_{23}^{21}\Gamma_{23}^{22} & \cdots & \Gamma_{23}^{2S} & \cdots & \Gamma_{2T}^{21}\Gamma_{2T}^{22} & \cdots & \Gamma_{2T}^{2S} \\ & \ddots & & & & & & & \\ \Gamma_{22}^{S1}\Gamma_{22} & \cdots & \Gamma_{22}^{SS}\Gamma_{23}^{S1}\Gamma_{23}^{S2} & \cdots & \Gamma_{23}^{SS} & \cdots & \Gamma_{2T}^{S1}\Gamma_{2T}^{S2} & \cdots & \Gamma_{2T}^{SS} \\ & & & \ddots & & & & & \\ & & & & \ddots & & & & \\ \Gamma_{T2}^{11}\Gamma_{T2}^{12} & \cdots & \Gamma_{T2}^{1S}\Gamma_{T3}^{11}\Gamma_{T3}^{12} & \cdots & \Gamma_{T3}^{1S} & \cdots & \Gamma_{TT}^{11}\Gamma_{TT}^{12} & \cdots & \Gamma_{TT}^{1S} \\ \Gamma_{T2}^{21}\Gamma_{T2}^{22} & \cdots & \Gamma_{T2}^{2S}\Gamma_{T3}^{21}\Gamma_{T3}^{22} & \cdots & \Gamma_{T3}^{2S} & \cdots & \Gamma_{TT}^{21}\Gamma_{TT}^{22} & \cdots & \Gamma_{TT}^{2S} \\ & & & & & & & \ddots & \\ \Gamma_{T2}^{S1}\Gamma_{T2}^{S2} & \cdots & \Gamma_{T2}^{SS}\Gamma_{T3}^{S1}\Gamma_{T3}^{S2} & \cdots & \Gamma_{T3}^{SS} & \cdots & \Gamma_{TT}^{S1}\Gamma_{TT}^{S2} & \cdots & \Gamma_{TT}^{SS} \end{pmatrix}$$

where we define

when $t_1 = t_2$ and $s_1 = s_2$

$$\Gamma_{t_1 t_2}^{s_1 s_2} = G_{t_1}(F_{t_1}^{-1}(p_{s_1}))[1 - G_{t_1}(F_{t_1}^{-1}(p_{s_1}))] + q_{s_1}(1-q_{s_1})$$

$$- 2\{G_{1t_1}(v_{s_1}, F_{t_1}^{-1}(p_{s_1})) - q_{s_1} G_{s_1}(F_{t_1}^{-1}(p_{s_1}))\}$$

when $t_1 = t_2$ and $s_1 \neq s_2$

$$\Gamma_{t_1 t_2}^{s_1 s_2} = G_{t_1} F_{t_1}^{-1}(p_{s^*}) - G_{1t_1}(v_2, F_{t_1}^{-1}(p_{s_1})) - G_{1t_1}(v_1, F_{t_1}^{-1}(p_{s_2}))$$

$$+ q_{s^*} - [G_{t_1} F_{t_1}^{-1}(p_{s_1}) - q_1][G_{t_1} F_{t_1}^{-1}(p_{s_2}) - q_2]$$

where $s^* = \min\{s_1, s_2\}$

when $t_1 \neq t_2$ and $s_1 = s_2$

$$\Gamma^{s_1 s_2}_{t_1 t_2} = G_{t_1 t_2}(F^{-1}_{t_1}(p_{s_1}), F^{-1}_{t_2}(p_{s_1})) - G_{1 t_1}(v_{s_1}, F^{-1}_{t_1}(p_{s_1}))$$

$$- G_{1 t_2}(v_{s_1}, F^{-1}_{t_2}(p_{s_1})) + q_{s_1}$$

$$- [G_{t_1}(F^{-1}_{t_1}(p_{s_1})) - q_{s_1}][G_{t_2}(F^{-1}_{t_2}(p_{s_1})) - q_{s_1}]$$

and when $t_1 \neq t_2$ and $s_1 \neq s_2$

$$\Gamma^{s_1 s_2}_{t_1 t_2} = G_{t_1 t_2}(F^{-1}_{t_1}(p_{s_1}), F^{-1}_{t_2}(p_{s_2}) - G_{1 t_1}(v_{s_2}, F^{-1}_{t_1}(p_{s_1}))$$

$$- G_{1 t_2}(v_{s_1}, F^{-1}_{t_2}(p_{s_2})) + q_{s*}$$

$$- [G_{t_1}(F^{-1}_{t_1}(p_{s_1})) - q_{s_1}][G_{t_2}(F^{-1}_{t_2}(p_{s_2})) - q_{s_2}].$$

Proof: Without loss of generality we prove the theorem for the case when T = 2 and S = 2.

Using a decomposition similar to that used in the proof of Theorem 1, we may write

$$\sqrt{m}\,\mathrm{vec}\begin{bmatrix} \hat{G}_2\hat{F}_2^{-1}(\hat{p}_1) - \hat{q}_1 - (G_2F_2^{-1}(p_1) - q_1) & \hat{G}_3\hat{F}_3^{-1}(\hat{p}_1) - \hat{q}_1 - (G_3F_3^{-1}(p_1) - q_1) \\ \hat{G}_2\hat{F}_2^{-1}(\hat{p}_2) - \hat{q}_2 - (G_2F_2^{-1}(p_2) - q_2) & \hat{G}_3\hat{F}_3^{-1}(\hat{p}_2) - \hat{q}_2 - (G_3F_3^{-1}(p_2) - q_2) \end{bmatrix}$$

$$= \sqrt{n}\underline{I} + \sqrt{m}\underline{II} + \sqrt{n}\underline{III}$$

where

$$\underline{I} = \begin{bmatrix} \hat{G}_2\hat{F}_2^{-1}(p_1) - \hat{q}_1 - (G_2F_2^{-1}(p_1)-q_1) \\ \hat{G}_2\hat{F}_2^{-1}(p_2) - \hat{q}_2 - (G_2F_2^{-1}(p_2)-q_2) \\ \hat{G}_3\hat{F}_3^{-1}(p_1) - \hat{q}_1 - (G_3F_3^{-1}(p_1)-q_1) \\ \hat{G}_3\hat{F}_3^{-1}(p_2) - \hat{q}_2 - (G_3F_3^{-1}(p_2)-q_2) \end{bmatrix}$$

$$\underline{II} = \begin{bmatrix} G_2\hat{F}_2^{-1}(\hat{p}_1) - G_2F_2^{-1}(p_1) \\ G_2\hat{F}_2^{-1}(\hat{p}_2) - G_2F_2^{-1}(p_2) \\ G_3\hat{F}_3^{-1}(\hat{p}_1) - G_3F_3^{-1}(p_1) \\ G_3\hat{F}_3^{-1}(\hat{p}_2) - G_3F_3^{-1}(p_2) \end{bmatrix}$$

and

$$\underline{III} = \begin{bmatrix} (\hat{G}_2\hat{F}_2^{-1}(\hat{p}_1)-G_2\hat{F}_2^{-1}(\hat{p}_1)) - (\hat{G}_2F_2^{-1}(p_1)-G_2F_2^{-1}(p_1)) \\ (\hat{G}_2\hat{F}_2^{-1}(\hat{p}_2)-G_2\hat{F}_2^{-1}(\hat{p}_2)) - (\hat{G}_2F_2^{-1}(p_2)-G_2F_2^{-1}(p_2)) \\ (\hat{G}_3\hat{F}_3^{-1}(\hat{p}_1)-G_3\hat{F}_3^{-1}(\hat{p}_1)) - (\hat{G}_3F_3^{-1}(p_1)-G_3F_3^{-1}(p_1)) \\ (\hat{G}_3\hat{F}_3^{-1}(\hat{p}_2)-G_3\hat{F}_3^{-1}(\hat{p}_2)) - (\hat{G}_3F_3^{-1}(p_2)-G_3F_3^{-1}(p_2)) \end{bmatrix} .$$

Obviously, given Theorem 1, it follows that $\sqrt{n}\underline{III} \xrightarrow{P} \underline{0}$. Thus, we need only show that

$$\sqrt{m}\underline{I} \xrightarrow{L} N_4(\underline{0},\underline{v}_1)$$

$$\sqrt{n}\underline{II} \xrightarrow{L} N_4(\underline{0},\underline{v}_2)$$

as $\underline{I}$ and $\underline{II}$ are independent.

$$\sqrt{n}\underline{I} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\underline{y}_j - \underline{v}_y)$$

where

$$\underline{y}_j = \begin{bmatrix} I(F_2^{-1}(p_1)-Y_{2j}) - I(v_1-Y_{1j}) \\ I(F_2^{-1}(p_2)-Y_{2j}) - I(v_2-Y_{1j}) \\ I(F_3^{-1}(p_1)-Y_{3j}) - I(v_1-Y_{1j}) \\ I(F_3^{-1}(p_2)-Y_{3j}) - I(v_2-Y_{1j}) \end{bmatrix}$$

and

$$\underline{v}_y = E(\underline{y}_j) = \begin{bmatrix} G_2F_2^{-1}(p_1) - q_1 \\ G_2F_2^{-1}(p_2) - q_2 \\ G_3F_3^{-1}(p_1) - q_1 \\ G_3F_3^{-1}(p_2) - q_2 \end{bmatrix}.$$

Let

$$\underline{\Sigma}^{Y} = \mathrm{var}(\underline{y}_j)$$

$$= \begin{bmatrix} \Gamma_{22}^{11} & \Gamma_{22}^{12} & \Gamma_{23}^{11} & \Gamma_{23}^{12} \\ & \Gamma_{22}^{22} & \Gamma_{23}^{21} & \Gamma_{23}^{22} \\ & " & \Gamma_{33}^{11} & \Gamma_{33}^{12} \\ & & & \Gamma_{33}^{22} \end{bmatrix}$$

where

$$\Gamma_{t_1t_2}^{s_1s_2} = \mathrm{cov}\{[I(F_{t_1}^{-1}(p_{s_1})-Y_{t_1j}) - I(v_{s_1}-Y_{1j})],[I(F_{t_2}^{-1}(p_{s_2})-Y_{t_2j})-I(v_{s_2}-Y_{1j})]\}.$$

When

$t_1 = t_2$ and $s_1 = s_2$, $\Gamma_{t_1t_2}^{s_1s_2}$ is given by Theorem 1,

when $t_1 \neq t_2$ and $s_1 = s_2$

$$\begin{aligned} \Gamma_{t_1t_2}^{s_1s_2} = {} & E\{I(F_2^{-1}(p_{s_1})-Y_{2j})I(F_3^{-1}(p_{s_1})-Y_{3j})\} \\ & - E\{I(F_2^{-1}(p_{s_1})-Y_{2j})I(v_{s_1}-Y_{1j})\} \\ & - E\{I(F_3^{-1}(p_{s_1})-Y_{3j})I(v_{s_1}-Y_{1j})\} \\ & + E\{I(v_{s_1}-Y_{1j})\} \\ & - E\{I(F_2^{-1}(p_{s_1})-Y_{2j})-I(v_{s_1}-Y_{1j})\}E\{I(F_3^{-1}(p_{s_1})-Y_{3j})-I(v_{s_1}-Y_{1j})\} \end{aligned}$$

$$= G_{23}(F_2^{-1}(p_{s_1}),F_3^{-1}(p_{s_1})) - G_{12}(v_{s_1},F_2^{-1}(p_{s_1}))$$

$$- G_{13}(v_{s_1},F_3^{-1}(p_{s_1})) + q_{s_1}$$

$$- [G_2(F_2^{-1}(p_{s_1}))-q_{s_1}][G_3(F_3^{-1}(p_{s_1})-q_{s_1}].$$

When $t_1 = t_2$ and $s_1 \neq s_2$

$$\Gamma_{t_1t_2}^{s_1s_2} = E\{I(F_{t_1}^{-1}(p_1)-Y_{t_1j})I(F_{t_1}^{-1}(p_2)-Y_{t_1j})\}$$

$$- E\{I(F_{t_1}^{-1}(p_1)-Y_{t_1j})I(v_2-Y_{1j})\}$$

$$- E\{I(v_1-Y_{1j})I(F_2^{-1}(p_2)-Y_{t_1j})\}$$

$$+ E\{I(v_1-Y_{1j})I(v_2-Y_{1j})\}$$

$$- E\{I(F_{t_1}^{-1}(p_1)-Y_{t_1j}) - I(v_1-Y_{1j})\}$$

$$\cdot E\{I(F_{t_1}^{-1}(p_2)-Y_{t_1j}) - I(v_2-Y_{1j})\}.$$

$p_1 < p_2 \Rightarrow F_{t_1}^{-1}(p_1) < F_{t_1}^{-1}(p_2)$

and so

$$\Gamma_{t_1t_2}^{s_1s_2} = G_{t_1}F_{t_1}^{-1}(p_{s*}) - G_{1t_1}(v_2,F_{t_1}^{-1}(p_1)) - G_{1t_1}(v_1,F_{t_1}^{-1}(p_2))$$

$$+ q_{s*} - [G_{t_1}F_{t_1}^{-1}(p_1)-q_1][G_{t_1}F_{t_1}^{-1}(p_2)-q_2],$$

where

$$s^* = \min\{s_1, s_2\}.$$

Finally,

when $t_1 \neq t_2$ and $s_1 \neq s_2$

$$\begin{aligned}
\Gamma^{s_1 s_2}_{t_1 t_2} &= E\{I(F_2^{-1}(p_{s_1}) - Y_{2j}) I(F_3^{-1}(p_{s_2}) - Y_{2j})\} \\
&\quad - E\{I(F_2^{-1}(p_{s_1}) - Y_{2j}) I(v_{s_2} - Y_{1j})\} \\
&\quad - E\{I(F_3^{-1}(p_{s_2}) - Y_{2j}) I(v_{s_1} - Y_{1j})\} \\
&\quad + E\{I(v_{s_1} - Y_{1j}) I(v_{s_2} - Y_{1j})\} \\
&\quad - E\{I(F_2^{-1}(p_{s_1}) - Y_{2j}) - I(v_{s_1} - Y_{1j})\} \\
&\quad \cdot E\{I(F_3^{-1}(p_{s_2}) - Y_{2j}) - I(v_{s_2} - Y_{1j})\} \\
&= G_{23}(F_2^{-1}(p_{s_1}), F_3^{-1}(p_{s_2})) - G_{12}(v_{s_2}, F_2^{-1}(p_{s_1})) \\
&\quad - G_{13}(v_{s_1}, F_3^{-1}(p_{s_2})) + q_{s^*} \\
&\quad - [G_2 F_2^{-1}(p_{s_1}) - q_{s_1}][G_3 F_3^{-1}(p_{s_2}) - q_{s_2}].
\end{aligned}$$

As the $\underline{y}_j$ are independent and identically distributed with mean $\underline{v}_y$ and variance-covariance matrix $\underline{\Sigma}^Y$ it follows that

$$\sqrt{n}\underline{I} = \frac{1}{\sqrt{n}}\left(\sum_{j=1}^{n}(\underline{y}_j - \underline{v}_y)\right) \xrightarrow{L} N_2(\underline{0}, \underline{\Sigma}^Y)$$

(see, for example, Anderson (1984)). Now,

$$\sqrt{m}\underline{II} = \sqrt{m}\begin{bmatrix} G_2\hat{F}_2^{-1}(\hat{p}_1) - G_2F_2^{-1}(p_1) \\ G_2\hat{F}_2^{-1}(\hat{p}_2) - G_2F_2^{-1}(p_2) \\ G_3\hat{F}_3^{-1}(\hat{p}_1) - G_3F_3^{-1}(p_1) \\ G_3\hat{F}_3^{-1}(\hat{p}_2) - G_3F_3^{-1}(p_2) \end{bmatrix}.$$

By Lemma 2 and application of the multivariate version of the delta method (see, for example, Bishop et al. (1975)) we have

$$\sqrt{m}\underline{II} \xrightarrow{L} N_2(\underline{0}, \underline{g}\,\underline{\Omega}^{-1}\,\underline{\Sigma}^X\underline{\Omega}^{-1}\underline{g})$$

where

$$\underline{g} = \text{diag}\begin{bmatrix} g_2(F_2^{-1}(p_1)) \\ g_2(F_2^{-1}(p_2)) \\ g_3(F_3^{-1}(p_1)) \\ g_3(F_3^{-1}(p_2)) \end{bmatrix}.$$

This completes the proof of Theorem 2.

3. FINITE SAMPLE PROPERTIES

3.1 Introduction

Monte Carlo studies were performed to examine the finite sampling behavior of some of the statistics developed in Chapter 2. Owing to the complexity in simulation, the Monte Carlo studies were confined to the case where a single diagnostic test based upon a continuous separator is to be compared with one based upon a discrete separator variable. Although most of the simulations were done with a binary separator for the discrete test some simulations were performed with the discrete test possessing two specificities.

An immediate consequence of Theorem 1 is that the standardized quantity

$$\frac{\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q} - [G_2F_2^{-1}(p)-q]}{\sqrt{V}}$$

is approximately distributed as N(0,1) where

$$V = \frac{q(1-q)}{n} + \frac{G_2F_2^{-1}(p)(1-G_2F_2^{-1}(p))}{n}$$

$$+ 2[\frac{d}{dp} G_2F_2^{-1}(p)]^2 \frac{[p-F(1,F_2^{-1}(p))]}{m}$$

$$- 2[\frac{G(1,F_2^{-1}(p) - qG_2(F_2^{-1}(p))}{n}] . \qquad (3.1.1)$$

Of course, in any finite-sample use V would have to be estimated and so we further extend the implication of Theorem 1 to be that

$$\frac{\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q} - [G_2F_2^{-1}(p)-q]}{\sqrt{\hat{V}}} \overset{\cdot}{\sim} N(0,1)$$

where we now define

$$\hat{V} = \frac{\hat{q}(1-\hat{q})}{n} + \frac{\hat{G}_2\hat{F}_2^{-1}(\hat{p})(1-\hat{G}_2\hat{F}_2^{-1}(\hat{p}))}{n}$$

$$+ 2\widehat{[\frac{d}{dp} G_2F_2^{-1}(p)]}^2 \frac{[\hat{p}-\hat{F}(1,\hat{F}_2^{-1}(\hat{p}))]}{m}$$

$$- 2[\frac{\hat{G}(1,\hat{F}_2^{-1}(\hat{p})) - \hat{q}\hat{G}_2(\hat{F}_2^{-1}(\hat{p}))}{n}]. \qquad (3.1.2)$$

This variance estimate is seen to be obtained by estimating each component of the asymptotic variance given by Theorem 1 with the "usual" sample estimators viz.; $\hat{q}$, $\hat{p}$ and $\hat{G}_2\hat{F}_2^{-1}(\hat{p})$ are the estimators described earlier, $\hat{F}(1,\hat{F}_2^{-1}(\hat{p}))$ and $\hat{G}(1,\hat{F}_2^{-1}(\hat{p}))$ are sample proportions estimating $\Pr\{X_1 = 1 \text{ and } X_2 \leq F_2^{-1}(p)\}$ and $\Pr\{Y_1 = 1 \text{ and } Y_2 \leq F_2^{-1}(p)\}$, respectively, and

$$\widehat{[\frac{d}{dp} G_2F_2^{-1}(p)]} = [\hat{G}_2\hat{F}_2^{-1}(\hat{p}+\nabla) - \hat{G}_2\hat{F}_2^{-1}(\hat{p}-\nabla)]/2\nabla,$$

is the "usual" naive estimator of

$$\frac{dG_2F^{-1}(p)}{dp} = \frac{g_2F_2^{-1}(p)}{f_2F_2^{-1}(p)} .$$

For $[\widehat{\frac{d}{dp} G_2F_2^{-1}(p)}]$ to be a consistent estimator, ∇ would have to converge slowly to zero which could make the derivative estimate unduly coarse at the sample sizes most likely used in practice. To obtain a "finer" estimate of the derivative, we have abandoned consistency and used $\nabla = \min(\hat{p}/2, \frac{1-\hat{p}}{2})$ for the estimator instead.

The Monte Carlo simulations were written in FORTRAN using IMSL subroutines to simulate sampling from various populations. In all cases, a random sample of m bivariate vectors

$$\underline{X}_i = \begin{pmatrix} X_{1i} \\ X_{2i} \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]$$

and a random sample of n bivariate vectors

$$\underline{Y}_j = \begin{pmatrix} Y_{1j} \\ Y_{2j} \end{pmatrix} \sim N_2\left[\begin{pmatrix} \mu_{Y_1} \\ \mu_{Y_2} \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right]$$

were generated. The discrete separator variable, X_1^* or Y_1^*, was then generated by setting

$$X_{1i}^* = \begin{cases} 1 & \text{if } X_{1i} \leq c^* \\ 0 & \text{otherwise} \end{cases}$$

and

$$Y^*_{1j} = \begin{cases} 1 & \text{if } Y_{1j} \leq c^* \\ 0 & \text{otherwise} \end{cases}$$

where c^* is the p^{th} percentile of the standard normal distribution and where μ_{Y_1} and σ_1 are selected so that $Pr(Y_{1j} \leq c^*) = q$. In this way the binary separator variable was formed which gave a diagnostic test having the desired specificity, p, and sensitivity, 1-q. The remaining parameters of the bivariate distributions, namely μ_{Y_2} and σ_2, were then selected so that the marginal distributions of X_2 and Y_2, F_2 and G_2 satisfied the condition that $1 - G_2F_2^{-1}(p) =$ preselected sensitivity of the continuous test. If the discrete test was to have two specificities, p_1 and p_2, and two sensitivities, $1-q_1$ and $1-q_2$, the above procedure involving the "discretization" of normally distributed random variables was suitably modified.

When the continuous separator was to have a uniform distribution the normally distributed random variables were transformed via the standard normal cdf. These variables were then modified to produce uniform separators which would yield a test of desired specificity, p, and sensitivity, (1-q). The exact transformations used were

$$X^*_{2i} = \Phi(X_{2i})$$

and

$$Y^*_{2j} = \Phi(Y_{2j}) + (p+(1-q)-1), \text{ where } \Phi = \text{standard normal cdf.}$$

Thus,

$$X^*_{2i} \sim U(0,1)$$

and

$$Y^*_{2j} \sim U([p-q],[p-q]+1)$$

and so, with cutoff c = p, we have

$$\text{Specificity} = \Pr(X^*_{2i} \leq p) = p$$

and

$$\text{Sensitivity} = \Pr(Y^*_{2j} \geq p) = 1-q.$$

In each simulation the correlation between tests measured upon the same individual was identical for individuals from the control and from the case populations. This was achieved by employing the same correlation in both bivariate normal populations, i.e., by setting $\rho = \sigma_{12}/(\sigma_1\sigma_2)$ in the previous parameterizations.

The correlation between the discretized version of a continuous random variable with another continuous random variable is not necessarily of the same exact magnitude, nor even in the same direction, as the correlation between the original two continuous variables. Since for purposes of simulation it was not considered important to obtain a specific numeric correlation between the discretized separator and its nondiscretized counterpart, an approximate "strength" of correlation between the two separator variables was attained by the following selection of the correlation, ρ, between the bivariate normal random variables which generated the separator variables:

Selected value of ρ	Resultant "strength" of correlation between discrete and continuous separators
.90	"Strong"
.50	"Moderate"
.10	"Weak"

The Monte Carlo simulations concentrated upon the behavior of a hypothesis testing procedure for the superiority of the continuous test. In particular, the Monte Carlo rejection rates of a nominal 0.05 size hypothesis test of

$$H_0: G_2F_2^{-1}(p)-q \geq 0 \quad \text{vs.} \quad H_A: G_2F_2^{-1}(p)-q < 0$$

were computed where the null hypothesis was rejected whenever the test statistic

$$\frac{\hat{G}_2\hat{F}_2^{-1}(\hat{p})-\hat{q}}{\sqrt{\hat{V}}}$$

was less than -1.645.

In a similar fashion, when the discrete test had two specificities, Monte Carlo rejection rates of tests of the null hypotheses

$$H_0: G_2F_2^{-1}(p_s)-q_s \geq 0, \; s = 1,2$$

and

$$H_0: \frac{1}{2} \sum_{s=1}^{2} (G_2 F_2^{-1}(p_s) - q_s) \geq 0$$

were computed.

3.2 Results of Monte Carlo Studies

Table 1 reports Monte Carlo rejection rates when the continuous separator is normally distributed for various sensitivities of the tests when the test based upon the discrete separator has specificity of 90% and the underlying separators are moderately correlated. All rejection rates are for the one-sided nominal 0.05 level hypothesis test described earlier in this chapter and are presented for common sample sizes of 30, 50 and 100. The rates for samples of size 30 or 50 are based upon 10,000 repetitions, while those for samples of size 100 are based upon 1,000 repetitions.

This table shows the one-sided hypothesis test to have good power in detecting the superiority of the diagnostic test based upon the continuous separator when this is the case. Furthermore, as would be hoped, the power of the procedure apparently increases as the common sample size increases or as the disparity in the sensitivities of the two tests increase.

The table does suggest, however, that the hypothesis testing procedure is anticonservative. All Monte Carlo estimates of the Type I error rate for the procedure exceed 0.05, the nominal level of the test. This consistent pattern, found not only here but in every other

Table 1. Monte Carlo rejection rates when continuous separator is normally distributed[a]

[1-q] Sensitivity of binary test	n=m	$[1 - G_2F_2^{-1}(p)]$ Sensitivity of continuous test 40%	60%	80%	90%
40%	30	0.097	0.438 [0.295][b]	0.834 [0.862]	0.951 [0.996]
	50	0.082	0.500 [0.414]	0.901 [0.970]	0.979 [≅1.0]
	100	0.059	0.635 [0.645]	0.987 [≅1.0]	0.997 [≅1.0]
60%	30		0.106	0.500 [0.396]	0.799 [0.835]
	50		0.080	0.589 [0.555]	0.885 [0.959]
	100		0.060	0.763 [0.810]	0.981 [0.999]
80%	30			0.093	0.332 [0.248]
	50			0.082	0.391 [0.345]
	100			0.064	0.522 [0.546]
90%	30				0.089
	50				0.068
	100				0.062

[a]Rejection rate of one-sided nominal 0.05 level hypothesis test when specificity = 0.90 and separator variables are moderately correlated. Rates for samples of size 30 or 50 based upon 10,000 repetitions. Rates for samples of size 100 based upon 1,000 repetitions.

[b]Theoretical power given in [ ].

simulation, gives credence to the conclusion that the procedure is intrinsically anticonservative. Additional support to this conclusion is obtained by noting that if the Type I error is 0.05 then the standard errors of the Monte Carlo Type I rejection rates is 0.002 when 10,000 repetitions are used, and 0.007 when 1,000 repetitions are used, and that most of the Monte Carlo estimates are well beyond three such standard errors of 0.05. Table 1 does show, nonetheless, the procedure to approach the nominal 0.05 level as sample size increases and that, generally, there is not a relationship between the degree the procedure is anticonservative and the level of the common sensitivities of the diagnostic tests.

Table 2 presents Monte Carlo studies of the effects of varying the specificity possessed by the test based upon the discrete separator. All rejection rates here are based upon 1,000 repetitions simulating the behavior of the one-sided hypothesis test described earlier when the continuous separator is normally distributed and the test results moderately correlated. Here again, the hypothesis testing procedure appears to be anticonservative and, in addition, this study suggests the anticonservativity worsens as the specificity of the discrete diagnostic test increases from 60% to 90%. Also, there is apparently a reduction in power as well when specificity is allowed to increase.

Table 3 reports the effects of varying the correlation between the separator variables upon the Monte Carlo rejection rates which, here, are based upon 1,000 repetitions each. Again, the rejection rates are for the earlier described one-sided hypothesis test when the continuous

Table 2. Effects of varying specificity on Monte Carlo rejection rates when continuous separator is normally distributed[a]

Type I Error:

[p] Specificity	n=m	Common sensitivity 40%	60%	80%	90%
60%	30	0.073	0.062	0.070	0.055
	50	0.063	0.070	0.054	0.046
	100	0.051	0.056	0.042	0.049
80%	30	0.087	0.063	0.073	0.062
	50	0.069	0.068	0.062	0.053
	100	0.073	0.053	0.050	0.055
90%	30	0.097	0.106	0.093	0.089
	50	0.082	0.080	0.082	0.068
	100	0.059	0.060	0.064	0.062

Power:

[p] Specificity	n=m	Discrete: Continuous:	Sensitivity 40% 90%	60% 90%	80% 90%
60%	30		0.990 [0.999][b]	0.877 [0.892]	0.320 [0.289]
	50		1.000 [≅1.0]	0.960 [0.981]	0.435 [0.405]
	100		1.000 [1.0]	0.999 [≅1.0]	0.615 [0.634]
80%	30		0.983 [0.998]	0.831 [0.874]	0.302 [0.274]
	50		1.000 [≅1.0]	0.948 [0.975]	0.380 [0.383]
	100		1.000 [1.0]	0.994 [≅1.0]	0.596 [0.603]
90%	30		0.951 [0.996]	0.799 [0.835]	0.332 [0.248]
	50		0.979 [≅1.0]	0.885 [0.959]	0.391 [0.345]
	100		0.997 [≅1.0]	0.981 [0.999]	0.522 [0.546]

[a]Rejection rate of one-sided nominal 0.05 level hypothesis test when separator variables are moderately correlated. Each rate based upon 1,000 repetitions.

[b]Theoretical power given in [ ].

Table 3. Effects of varying the correlation between separator variables on Monte Carlo rejection rates when the continuous separator is normally distributed[a]

Sensitivity			Correlation		
Discrete	Continuous	n=m	Weak	Moderate	Strong
Type I Error:					
40%	40%	30	0.075	0.097	0.164
		50	0.081	0.082	0.100
		100	0.061	0.059	0.081
60%	60%	30	0.098	0.106	0.150
		50	0.085	0.080	0.099
		100	0.072	0.060	0.074
80%	80%	30	0.087	0.093	0.117
		50	0.082	0.082	0.087
		100	0.070	0.064	0.070
90%	90%	30	0.083	0.089	0.055
		50	0.080	0.068	0.071
		100	0.047	0.062	0.043
Power:					
60%	80%	30	0.437 [0.332][b]	0.500 [0.396]	0.758 [0.601]
		50	0.512 [0.467]	0.589 [0.553]	0.845 [0.791]
		100	0.685 [0.714]	0.763 [0.810]	0.953 [0.966]
60%	90%	30	0.718 [0.759]	0.799 [0.835]	0.965 [0.933]
		50	0.846 [0.917]	0.885 [0.951]	0.985 [0.992]
		100	0.957 [0.996]	0.981 [0.999]	1.00 [≅1.0]
80%	90%	30	0.287 [0.214]	0.332 [0.248]	0.486 [0.376]
		50	0.295 [0.293]	0.391 [0.345]	0.576 [0.529]
		100	0.456 [0.465]	0.522 [0.546]	0.789 [0.783]

[a] Rejection rate of one-sided nominal 0.05 level hypothesis test when specificity = 0.90. Each rate based upon 1,000 repetitions.

[b] Theoretical power given in [ ].

separator is normally distributed and when the discrete test has specificity 90%.

In many cases the anticonservativity of the procedure worsened as correlation strengthened. However, there are several exceptions to this pattern, and no distinct trend is apparent. Power, on the other hand, shows itself to be consistently increasing as the correlation between separator variables increases. This effect is not surprising if we recall the variance expression

$$V = \frac{q(1-q)}{n} + \frac{G_2F_2^{-1}(p)(1-G_2F_2^{-1}(p))}{n}$$

$$+ 2[\frac{d}{dp} G_2F_2^{-1}(p)]^2 \frac{[p-F(1,F_2^{-1}(p))]}{m}$$

$$- \frac{2[G(1,F_2^{-1}(p)) - qG_2F_2^{-1}(p)]}{n} .$$

Now $\Pr\{X_2 \leq F_2^{-1}(p)/X_1 = 1\} = \int_{-\infty}^{c}\int_{-\infty}^{F_2^{-1}(p)} f_{X_2/X_1}^{(w_2)} f_{X_1}^{(w_1)} dw_2 dw_1 / P(X_1 = 1)$

where f_{X_1} represents the N(0,1) density and f_{X_2/X_1} represents the density of X_2 given X_1, namely $N(\mu_2 + \rho X_1, (1-\rho^2))$. Since as ρ increases the conditional density of X_2 given X_1 becomes more concentrated about the upward sloping line $\mu_2 + \rho X_1$, it follows that the double integral of the above expression increases in ρ. Thus, when (X_1, X_2) is bivariate normal, $[p-F(1,F_2^{-1}(p))] = [p - p \Pr\{X_2 \leq F_2^{-1}(p)/X_1 = 1\}]$ is a decreasing function of ρ. Likewise, $-2[G(1,F_2^{-1}(p)) - qG_2F_2^{-1}(p)] =$

$2q[G_2F_2^{-1}(p) - Pr\{Y_2 \leq F_2^{-1}(p)/Y_1 = 1\}]$ is a decreasing function of ρ. And so, the effect of increasing "correlation" between the tests is to reduce the variance of the estimated difference in sensitivities. This gives a more powerful procedure.

Table 4 explores the impact of an alternative distribution for the continuous separator variable. Here, the rejection rates are reported for the one-sided hypothesis testing procedure when the continuous separator is uniformly distributed and is moderately correlated with the discrete separator. The specificity of the discrete test has been set at 90% and all rejection rates are based on 1,000 repetitions.

Comparing these results to those given in Table 1 show the behavior of the procedure to be fairly "robust" to the different distributions of the continuous separator variable. The anticonservativity problem is suggested here as well, although it does seem to be a bit less severe in certain cases. However, the problem does appear here to worsen with increasing common sensitivity. Power, on the other hand, appears to be consistently better when the continuous distribution is uniform.

Finally, Tables 5 and 6 report Monte Carlo rejection rates when the discrete test has two specificities and the continuous separator is normally distributed. Rejection rates when testing each of the following null hypotheses,

$$H_0: G_2F_2^{-1}(p_s) - q_s \geq 0, \quad s = 1,2,$$

and

Table 4. Monte Carlo rejection rates when continuous separator is uniformly distributed[a]

[1-q] Sensitivity of binary test	n=m	$[1 - G_2F_2^{-1}(p)]$ Sensitivity of continuous test 40%	60%	80%	90%
40%	30	0.086	0.570 [0.486][b]	0.963 [0.954]	0.997 [0.996]
	50	0.076	0.723 [0.668]	0.995 [0.996]	0.999 [≅1.0]
	100	0.059	0.913 [0.902]	1.000 [≅1.0]	1.000 [1.0]
60%	30		0.099	0.609 [0.519]	0.908 [0.835]
	50		0.061	0.749 [0.705]	0.966 [0.959]
	100		0.058	0.933 [0.926]	0.999 [0.999]
80%	30			0.106	0.384 [0.248]
	50			0.071	0.457 [0.345]
	100			0.063	0.618 [0.546]
90%	30				0.118
	50				0.116
	100				0.103

[a]Rejection rate of one-sided nominal 0.05 level hypothesis test when specificity = 0.90 and when separator variables are moderately correlated. Rates are based upon 1,000 repetitions.

[b]Theoretical power given in [ ].

Table 5. Monte Carlo estimates of Type I error rate when discrete test has two specificities and continuous separator is normally distributed[a]

Comparison at specificity	Common sensitivity	Monte Carlo type I error rate
80%	60%	0.084
90%	40%	0.083
combined[b]		0.082
80%	80%	0.069
90%	40%	0.104
combined		0.089
80%	90%	0.075
90%	40%	0.109
combined		0.097
80%	80%	0.069
90%	60%	0.093
combined		0.087
80%	90%	0.060
90%	60%	0.118
combined		0.101
80%	90%	0.050
90%	80%	0.066
combined		0.058

[a]Rejection rate of one-sided nominal 0.05 level hypothesis test when test results are moderately correlated and common sample size is 50. Rates based upon 1,000 repetitions.

[b]Average of the comparisons made at the previous two specificities.

Table 6. Monte Carlo estimates of power when discrete test has two specificities and continuous separator is normally distributed[a]

Comparison at specificity	Sensitivity of discrete test	Sensitivity of continuous test	Monte Carlo estimate of power
80%	60%	80%	0.242
90%	40%	40%	0.127
combined[b]			0.166
80%	60%	90%	0.330
90%	40%	40%	0.140
combined			0.219
80%	60%	80%	0.498
90%	40%	60%	0.413
combined			0.506
80%	60%	90%	0.946
90%	40%	80%	0.901
combined			0.953
80%	80%	90%	0.520
90%	60%	80%	0.617
combined			0.652

[a]Rejection rate of one-sided nominal 0.05 level hypothesis test when test results are moderately correlated and common sample size is 50. Rates based upon 1,000 repetitions.

[b]Average of the comparisons made at the previous two specificities.

$$H_0: \frac{1}{2} \sum_{s=1}^{2} G_2 F_2^{-1}(p_s) - q_s \geq 0$$

where $p_1 = .80$ and $p_2 = .90$, are given when the null hypothesis is true (Table 5) and for various alternatives (Table 6). In all cases, the simulations were done with the separator variables being moderately correlated with common sample size of 50 in 1,000 repetitions.

Further confirmation of anticonservativity in the procedure, both in its "single-specificity" form and "multiple-specificity" extension,[6] is provided in Table 5. Consistent with the results from Table 2, the anticonservancy of the procedure at the higher specificity (90%) is worse than at the lower specificity (80%). Perhaps not too surprisingly, the anticonservativity of the combined procedure lies between the previous two values. In Table 6, power is likewise shown to decrease with increasing specificity for each of the "single-specificity" comparisons and to be between these values when the comparisons are averaged.

3.3 Investigations into the Anticonservativity of the Statistic

Table 7 presents Monte Carlo studies of the behavior of the standardized statistic and two alternative methods of standardization. The standardized statistic is, of course,

---

[6] The "multiple-specificity" extension referred to here is the simple average of the comparisons at each specificity available to the discrete test.

Table 7. Monte Carlo performance of various standardizations of the test statistic when continuous separator is normally distributed[a]

n=m	Rejection rates[b]			Mean			Variance		
	1	2	3	1	2	3	1	2	3
30	0.098	0.044	0.040	-0.14	-0.09	-0.01	1.60	0.90	1.37
50	0.078	0.035	0.036	-0.10	-0.06	≅0	1.40	0.94	1.21
100	0.072	0.047	0.027	-0.09	-0.04	-0.03	1.20	1.01	0.98

Standardizations:

1 = as developed in thesis,

2 = standardized using true values of the variance components,

3 = standardized using true value of the density ratio in estimating the variance.

[a]Simulations done with specificity = 0.90, common sensitivity = 0.60 and the separator variables moderately correlated. Rates based upon 1,000 repetitions.

[b]Based upon a one-sided nominal 0.05 level hypothesis test.

$$[\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q}]/\sqrt{\hat{V}}$$ (referred to as standardization #1 in the table),

where $\hat{V}$ is as defined in (3.1.2). The two other standardizations are

$$[\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q}]/\sqrt{V}$$ (standardization #2)

which is seen to be standardization of the observed statistic by the asymptotic variance term (3.1.1) (considered to be known) and,

$$[\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q}]/\sqrt{V^*}$$ (standardization #3)

where

$$V^* = \frac{\hat{q}(1-\hat{q})}{n} + \frac{\hat{G}_2\hat{F}_2^{-1}(\hat{p})(1-\hat{G}_2\hat{F}_2^{-1}(\hat{p}))}{n}$$

$$+ 2\left[\frac{d}{dp} G_2F_2^{-1}(p)\right]^2 \frac{[\hat{p}-\hat{F}(1,\hat{F}_2^{-1}(\hat{p}))]}{m}$$

$$- 2\left[\frac{\hat{G}(1,\hat{F}_2^{-1}(\hat{p}))-\hat{q}\hat{G}_2(\hat{F}_2^{-1}(\hat{p}))}{n}\right] .$$

Note that V^* differs from $\hat{V}$ only in that V^* contains the true value of the derivative $\frac{d}{dp} G_2F_2^{-1}(p)$.

Table 7 shows the alternate standardizations to behave quite well. Neither suffer from anticonservancy and both more closely approximate a N(0,1) type random variable than does standardization #1. The fact that such improvement can be realized by the alternate standardizations suggests the origins of the anticonservativity problem lie partially in the estimation of the variance of the statistic $\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q}$. That substantial improvement is afforded by the third standardization method suggests, further, that much of the problem perhaps lies with the derivative estimation.

However, examination of the Monte Carlo performance of the estimators of the components of the asymptotic variance, including that of the derivative, showed they each performed well in estimating their respective component <u>given</u> a fixed value for the estimated specificity. In particular, for fixed $\hat{p}$, the mean values of the variance component

estimators $\hat{G}_2\hat{F}_2^{-1}(\hat{p})$, $\widehat{\frac{d}{dp}\, G_2F_2^{-1}(\hat{p})}$, $\hat{F}(1,\hat{F}_2^{-1}(\hat{p}))$ and $\hat{G}(1,\hat{F}_2^{-1}(\hat{p}))$ were close to $G_2F_2^{-1}(\hat{p})$, $\frac{d}{dp}\, G_2F_2^{-1}(\hat{p})$, $F(1,F_2^{-1}(\hat{p}))$ and $G(1,F_2^{-1}(\hat{p}))$, respectively. Thus, the variance estimator $\hat{V}$ worked well in estimating the asymptotic variance when specificity was "known" to be a particular $\hat{p}$.

The anticonservativity problem apparently arises then mainly from the variability in the estimation of the specificity of the discrete test. Of course, such estimation induces variability in the entire estimation process, but extra variability alone does not account for anticonservativity. More importantly, as is shown below, small changes in the specificity estimate can make substantial changes in the standardized statistic thus inflating rejection rates.

First, as $G_2F_2^{-1}(\hat{p})$ is an increasing function of $\hat{p}$ and as the Monte Carlo studies show $\hat{G}_2\hat{F}_2^{-1}(\hat{p})$ to be a "good" estimator of $G_2F_2^{-1}(\hat{p})$ when $\hat{p}$ is considered fixed, the average value of $\hat{G}_2\hat{F}_2^{-1}(\hat{p})$ is seen to also be an increasing function of $\hat{p}$. This implies that the estimated difference $\hat{G}_2\hat{F}_2^{-1}(\hat{p}) - \hat{q}$ tends, on the average, to underestimate the true difference as $\hat{p}$ decreases, thus making it more likely the standardized statistic will fall below the critical -1.645 in the one-sided hypothesis test used in the Monte Carlo studies.

Second, note that when G_2 and F_2 are normal cdfs, with F_2 being the standard normal cdf,

$$\frac{d}{dp} G_2 F_2^{-1}(p) = \frac{g_2 F_2^{-1}(p)}{f_2 F_2^{-1}(p)}$$

$$= \frac{1}{\sigma_{Y_2}^2} \exp\left\{ - \frac{1}{2}\left(\frac{F_2^{-1}(p) - \mu_{Y_2}}{\sigma_{Y_2}} \right)^2 + \frac{1}{2}(F_2^{-1}(p))^2 \right\},$$

and when, as in the simulations, $\mu_{Y_2} > 0$ and $\sigma_{Y_2} \geq 1$, this derivative is, therefore, an increasing function of p. In addition, note that, as Table 8 shows, the derivative can be a very sharply increasing function of p.

Recalling equation (3.1.2), we observe that $\hat{V}$ is an increasing function of the derivative estimate. And so, when $\hat{p}$ underestimates p, $\widehat{\frac{d}{dp} G_2 F_2^{-1}(p)}$ being a "good" estimator will estimate $\frac{d}{dp} G_2 F_2^{-1}(\hat{p})$, which can be appreciably smaller than $\frac{d}{dp} G_2 F_2^{-1}(p)$. This will reduce $\hat{V}$ which will, in turn, tend to inflate the value of the standardized statistic, leading to more frequent rejections of the null hypothesis.

In sum, as $\hat{p}$ underestimates p, the numerator of the standardized statistic tends to be inflated and the denominator tends to be deflated thereby increasing the likelihood the standardized statistic will achieve an extreme value. This greater likelihood of extreme values manifests itself in the observed anticonservancy of the one-sided hypothesis testing procedure.

Monte Carlo studies of other hypothesis testing procedures, both one-sided at sizes different than 0.05 and two-sided procedures, demonstrate anticonservativity and lend credence to the above explanation of

the problem of anticonservativity. For example, as is shown in Table 9, in simulating the performance of a two-sided hypothesis test the Monte Carlo Type I rejection rates for lower-tail values are much more inflated than those observed for the upper tail (each should be 0.05). This demonstrates the inflationary effects of underestimating p and "dampening" effects of overestimating p.

Table 8. Various values of $dG_2F_2^{-1}(p)/dp$ when $\mu_{Y_2} = 2.56$ and $\sigma^2_{Y_2} = 1$[a]

p	$dG_2F_2^{-1}(p)/dp$
.6	0.0023
.8	0.0793
.9	1.0000
.95	8.5901

[a]The parameterization of the distribution of Y_2 here was used in Monte Carlo studies when specificity = .90 and the tests had common sensitivity = .90.

Since $\hat{p}$ is a consistent estimator of p, the obvious solution to the anticonservativity problem is to take very large samples, expecially of controls. In cases where extremely large samples cannot be employed, several modifications to the standardized statistic are suggested by the discussion above.

Table 9. Lower- and upper-tail Type I Monte Carlo rejection rates

Common sensitivity	Rejection rates[a] Lower tail	Upper tail
.40	.082	.060
.60	.080	.057
.80	.082	.046
.90	.068	.037

[a]Rates based on 1,000 repetitions of a two-sided hypothesis test at nominal size 0.10 when specificity of discrete test is 90%, continuous separator is normally distributed and test results are moderately correlated. Common sample size of 50 was used.

First, we might consider adding $1/m$ to $\hat{p}$ in the numerator of the standardized statistic. Although this will bias the estimate of p, it will nonetheless act to dampen the likelihood of severely underestimating a null difference in sensitivities between the two tests. Note that this modification requires no additional assumptions be placed upon F_2 or G_2 for it to be an appropriate "fix" to anticonservancy.

Second, if we should assume that, as in the case when F_2 and G_2 are Normal cdfs, the derivative $\frac{d}{dp} G_2 F_2^{-1}(p)$ is an increasing function of p, we might add $1/m$ to $\hat{p}$ in the denominator of the standardized statistic. Doing this lessens the likelihood of extreme values of the standardized statistic while preserving the unbiasedness of the numerator.

Third, we might simply choose to overestimate p with, for example, $\hat{p} + 1/m$. Thus, we would use $\hat{p} + 1/m$ in both numerator and denominator

of the standardized statistic. Again, as a "fix" to anticonservativity this method makes sense only insofar as the derivative is an increasing function of p.

Finally, we might envision a procedure which would form an interval estimate of p around $\hat{p}$ and then take the largest variance estimate arising from specificity values within that interval. This procedure would act to reduce the rejection rate by using a "likely-inflated" variance estimate. Assuming monotonicity in the variance as a function of $\hat{p}$ across this interval, and using the interval estimate $\hat{p} \pm 1/m$, leads to using the largest of $\{\hat{V}(\hat{p} - 1/m), \hat{V}(\hat{p}), \hat{V}(\hat{p} + 1/m)$ in the denominator of the standardized statistic, where $V(\cdot)$ refers to the variance estimate computed with $(\cdot)$ as the value of the estimate of p.

Tables 10 and 11 report Monte Carlo estimates of, respectively, Type I Error and Power for the previous four modifications to the standardized test statistic when the continuous separator is normally distributed, the separator variables moderately correlated, and when the test based upon the discrete separator has specificity of 90%. The estimates are based upon 1,000 repetitions simulating the behavior of a one-sided nominal 0.05 level hypothesis test.

As is to be expected, the Type I Error rates of the four modifications are less than that of the unmodified standardized test statistic. Of course, as is also to be expected, the power of the four modifications are apparently less than that of the unmodified statistic.

Overall, no single modification yet considered was always conservative. The best "fix" to anticonservativity comes from adding 1/m to the

Table 10. Monte Carlo estimates of Type I Error rates for modifications of the test statistic when continuous separator is normally distributed[a]

Common sensitivity	Common sample size	Modification 1	2	3	4	5
40%	30	0.101	0.060	0.078	0.076	0.070
	50	0.087	0.053	0.067	0.064	0.064
	100	0.064	0.036	0.052	0.043	0.050
60%	30	0.088	0.037	0.057	0.059	0.055
	50	0.091	0.046	0.075	0.065	0.072
	100	0.073	0.045	0.060	0.057	0.057
80%	30	0.076	0.032	0.045	0.049	0.044
	50	0.075	0.037	0.051	0.053	0.048
	100	0.063	0.034	0.045	0.046	0.044
90%	30	0.050	0.027	0.035	0.039	0.033
	50	0.057	0.033	0.040	0.046	0.040
	100	0.054	0.035	0.039	0.049	0.039

Modification codes:

1 = unmodified standardized statistic,

2 = $\frac{1}{m}$ added to specificity estimate in both numerator and denominator of standardized statistic,

3 = $\frac{1}{m}$ added only in denominator of standardized statistic,

4 = $\frac{1}{m}$ added to specificity estimate only in numerator of standardized statistic,

5 = use maximum variance estimate in denominator.

[a]Rejection rate of one-sided nominal 0.05 level hypothesis test when specificity = 90% and separator variables are moderately correlated. Rates based upon 1,000 repetitions.

Table 11. Monte Carlo estimates of power for modifications of the test statistic when continuous separator is normally distributed[a]

Sensitivity of discrete test	Sensitivity of continuous test	Common sample size	Modification					Theoretical power
			1	2	3	4	5	
.60	.80	30	0.491	0.325	0.386	0.402	0.362	0.396
		50	0.573	0.424	0.502	0.494	0.488	0.555
		100	0.771	0.674	0.721	0.713	0.715	0.810
	.90	30	0.804	0.661	0.739	0.728	0.716	0.835
		50	0.901	0.812	0.855	0.869	0.844	0.959
		100	0.984	0.964	0.972	0.975	0.971	0.999
.80	.90	30	0.274	0.173	0.201	0.220	0.188	0.248
		50	0.340	0.241	0.272	0.295	0.264	0.345
		100	0.527	0.428	0.472	0.466	0.470	0.546

Modification codes:

1 = unmodified standardized statistic,

2 = $\frac{1}{m}$ added to specificity estimate in both numerator and denominator of standardized statistic,

3 = $\frac{1}{m}$ added to specificity estimate only in denominator of standardized statistic,

4 = $\frac{1}{m}$ added to specificity estimate only in numerator of standardized statistic,

5 = use maximum variance estimate in denominator.

[a]Rejection rate of one-sided nominal 0.05 level hypothesis test when specificity = 0.90 and separator variables are moderately correlated. Rates based upon 1,000 repetitions.

estimate of p in both the numerator and denominator of the standardized statistic. However, even this procedure was apparently slightly anti-conservative several times. Perhaps the most defensible procedure is the one which uses the maximum variance estimate in the denominator as this procedure requires no additional assumptions in order to be considered as a reasonable "fix" and retains the unbiasedness in estimating the true difference in sensitivities. This procedure compared favorably as a "fix" with the other modifications considered thus far, only on five of 12 simulations exhibiting any sort of anticonservativity, and was always less anticonservative than the unmodified standardized statistic. Also, the power of the procedure compared favorably with that of the unmodified standardized statistic, but in all cases was slightly less powerful.

A final avenue explored for the correction of the anticonservativity problem was the use of the transformation

$$\gamma = \ln\left[\frac{1-G_2F_2^{-1}(p)}{1-q}\right].$$

It was hoped this transformation would quicken the approach to normality taken by the associated test statistic

$$\ln\left[\frac{1-\hat{G}_2\hat{F}_2^{-1}(\hat{p})}{1-\hat{q}}\right] / \hat{V}_\gamma$$

where $\hat{V}_\gamma$ is the sample estimate of the finite sample variance. This variance is, via Taylor series expansion, approximately equal to

$$\left[\frac{1}{1-q}\right]^2 \frac{q(1-q)}{n}$$

$$+\left[\frac{1}{1-G_2F_2^{-1}(p)}\right]^2\left\{\frac{G_2F_2^{-1}(p)(1-G_2F_2^{-1}(p))}{n}+\frac{\left[\frac{d}{dp}\,G_2F_2^{-1}(p)\right]^2[p-F(1,F_2^{-1}(p))]}{m}\right\}$$

$$-2\left\{\frac{G(1,F_2^{-1}(p))-qG_2F_2^{-1}(p)}{n}-\frac{1}{2}\,\frac{\left[\frac{d}{dp}\,G_2F_2^{-1}(p)\right]^2[p-F(1,F_2^{-1}(p))]}{m}\right\}.$$

As shown in Table 12, the procedure based on γ is apparently as anticonservative as the untransformed statistic. It is reasonable to expect that here, as before, the anticonservativity arises largely due to the necessity of the estimation of the specificity of the diagnostic test based upon the discrete separator.

Table 12. Comparison of Monte Carlo Type I Error Rates for the transformation against the untransformed statistic

Common sensitivity	Common sample size	Type I Error Rates[a]	
		Transformed	Untransformed
60%	30	0.090	0.101
	100	0.080	0.070
80%	30	0.070	0.086
	100	0.067	0.067
90%	30	0.034	0.042
	100	0.049	0.054

[a]Rates based upon one-sided nominal 0.05 level hypothesis test for superiority of test based upon continuous separator. Specificity of test based upon discrete separator = 90%, separators moderately correlated.

4. APPLICATIONS

4.1 Introduction

In this chapter, several examples of the application of the statistic are presented. The first example presents application of the statistic to the comparison of two screening procedures for gastrointestinal malignancies, one of which is based upon a binary response variable and the other is to be based upon the result of a continuous assay measurement. The other example demonstrates application of the statistic to the comparison of two tests based upon continuous variables with a third test based upon a separator variable having 13 possible ordered outcomes. The aim of the investigator who collected the data for this example was to compare the three procedures for their ability to identify patients who would benefit from a certain surgical procedure.

4.2 An Example of the Comparison of a Test Based Upon a Binary Separator with One Based Upon a Continuous Separator

In 1983, the paper "A review of American Cancer Society estimates of cancer cases and deaths" (Silverberg and Lubera, 1983) reported that gastrointestinal malignancies accounted for nearly one third of all cancer deaths in the United States. Since early detection is crucial in successful treatment of this disease, an important medical problem is the selection of a "best" screen for gastrointestinal malignancies.

Measurement of fecal hemoglobin has been identified as a possible "marker" for the presence of gastrointestinal malignancies (Ahlquist et al. (1985)). Presently there are, however, but two methods for the

detection of elevated fecal blood which are noninvasive, inexpensive, and which may be done quickly enough to lend themselves to be used for screening large numbers of individuals.

One of these procedures, Hemoccult, is a pad test which results in either a negative (fecal blood level not elevated) or positive reading. Hemoccult needs no special equipment and has been used in practice for several years. The other procedure, HemoQuant, is an assay of the miligrams of blood per gram of stool. This procedure, unlike other quantitative measures of fecal blood levels, is easily, quickly and inexpensively done in the physician's own laboratory.

Ahlquist et al. (1985) compared the sensitivities of the previous two procedures for detecting elevated fecal blood levels. Both procedures were applied to each individual enrolled in the study. The following example is based upon this study and incorporates the data from m = 522 healthy volunteers or patients admitted to the hospital with gastrointestinal symptoms but with normal gastrointestinal studies, and from n = 222 patients drawn consecutively from a population of patients who had diagnostic studies indicating the presence of gastrointestinal lesions. All such examinations were performed without knowledge of fecal test results.

Hemoccult had an empirical specificity of 96.6%. The associated 96.6^{th} percentile of the HemoQuant assay of fecal blood was, for the controls, estimated to be the value 4.56mg hemoglobin/gram stool. Using 4.56 as the cutoff value when classifying individuals with the

HemoQuant separator variable, a sensitivity of 26.68% was observed. The estimated sensitivity of the test based upon the binary separator (i.e., Hemoccult) was 24.22%.

The difference in sensitivities between the two screens is, therefore, estimated to be

$$[1-\hat{q}] - [1-\hat{G}_2\hat{F}_2^{-1}(\hat{p})] = -1.46\%$$

having an estimated standard error of 5.4%. The standardized statistic is thus 0.2703 which, under normal theory, would possess a p-value of approximately 0.39 for testing

H_0: Sensitivity of Hemoccult is at least as great as the sensitivity of HemoQuant.

Note that an approximate 95% confidence interval on the true difference in sensitivities when both tests have the same specificity is the interval (- 12.04%, 9.12%).

In light of the anticonservancy shown in the Monte Carlo studies and the large sample sizes employed here, we may conclude that any difference in sensitivities between the two tests is neglible when they both possess 96.6% specificity.

4.3 An Example Comparing a Test Based Upon a S-Nary Separator Variable with Two Tests Each Based Upon Continuous Separators

As reported in Delong et al. (1988), a dilema concerning the management of patients known to have ovarian carcinoma is in determining

whether surgical correction of intestinal obstruction will benefit the patient. Some authors propose that patients who survive longer than two months postoperatively be declared to have "benefited" from the surgery. A preoperative scoring system based upon this criterion has been devised for use as a screening test in determining if a patient will benefit from surgery. This scoring system, known as the Krebs-Goplerud score (herein referred to as the K-G score), assigns one of 13 possible integral values between zero and 12 to a patient, with higher scores being associated with patients who benefit from the surgery.

The Delong et al. (1988) paper compares the K-G scoring system against two other preoperatively measured indices, albumin and total protein, as possible screens for patients likely to benefit from surgery. Each of these latter two indices are continuous measurements which are positively associated with the patient's nutritional status, and thus it is reasonable to expect them to be positively associated with increased likelihood of benefit from the surgery. These three screening procedures are compared by Delong et al. using a procedure developed in their paper for the nonparametric comparison of the areas under correlated ROC curves.

The data analyzed by the Delong method are also amenable to analysis by the methods developed in my dissertation. Here we desire the comparison of a screening procedure based upon a discrete separator variable, the K-G score, with two screening procedures each of which are to be based upon continuous separators. Following the method developed in my thesis, each of the S=13 possible specificities of the screening procedure

based upon the discrete separator will be estimated (each possible value of the K-G score is a candidate cutoff for this screen). Then, "stepping through" each such estimated specificity, cutoffs for each continuous test will be determined which give common specificity to the tests, and then sensitivities of the three tests at this estimated specificity will be estimated and compared. In addition, estimates of the variances and covariances of these comparisons of sensitivity will be made.

The following estimates are based upon measurements of albumin, total protein and K-G score made prior to corrective surgery for abdominal obstruction for 49 consecutively-entered ovarian cancer patients at Duke University Medical Center. Using the criterion mentioned earlier, 37 of these patients survived more than two months postoperatively and are considered surgical successes and the remaining 12 are considered surgical failures. Six of the successes had missing data and are excluded from analysis: five of these patients had no measurements for either albumin or total protein, and one lacked measurement of total protein. Thus, considering a "case" to be a surgical success and a "control" to be a surgical failure, we have n=31 and m=12.

Table 13 reports estimated differences in sensitivities between the diagnostic procedure based upon the K-G score and the procedure based upon albumin, and standard errors of these estimates, at each of the specificities estimated to be available to the screen based upon the K-G score. The table presents similar statistics for the comparison of sensitivities between the K-G based test and the test based upon total

protein. In each case, the estimated difference is taken with the sensitivity of the test based upon the discrete score as the minuend.

Table 13. Comparison of Sensitivities at each estimated Specificity of the test based upon Krebs-Goplerud (K-G) scoring system

K-G cutoff	Est. spec. (%)	(K-G) - Albumin			(K-G) - Total Protein		
		Est. dif. (%)	SE(%)	99.8% C.I.[a]	Est. dif. (%)	SE(%)	99.8% C.I.[a]
0	0	0	0	.	0	0	.
1	0	0	0	.	0	0	.
2	8.3	0	2.24	[-6.5, 6.5]	0	0	.
3	16.6	3.2	10.20	[-26.2,32.6]	0	6.48	[-18.7,18.7]
4	41.7	12.9	12.41	[-22.8,48.6]	29.0	18.08	[-23.1,81.1]
5	50.0	0	11.66	[-33.6,33.6]	16.1	21.93	[-47.1,79.3]
6	75.0	-6.5	25.38	[-79.6,66.6]	22.6	18.65	[-31.1,76.3]
7	83.0	-35.5	31.54	[-126.3,55.3]	-16.1	12.37	[-51.7,19.5]
8	100	-6.5	7.81	[-29.0,16.0]	0	6.48	[-18.7,18.7]
9	100	-12.9	6.00	[-30.2, 4.4]	-6.5	4.36	[-19.1, 6.1]
10	100	-12.9	6.00	[-30.2, 4.4]	-6.5	4.36	[-19.1, 6.1]
11	100	-12.9	6.00	[-30.2, 4.4]	-6.5	4.36	[-19.1, 6.1]
12	100	-12.9	6.00	[-30.2, 4.4]	-6.5	4.36	[-19.1, 6.1]

[a]99.8% confidence interval upon difference in sensitivities chosen to give 90% simultaneous confidence coverage.

Included with the estimated difference in sensitivities are 99.8% confidence intervals upon the true difference. These confidence intervals are based upon the standard normal distribution and utilize 2.88(SE) as the interval half-width. The level of confidence for each interval was

chosen via the Bonferroni method to ensure 90% confidence coverage by the twenty-six intervals.

All of the twenty-one confidence intervals placed when the SE of the estimated difference was not equal to zero contain zero and, therefore, simultaneously show the K-G scoring procedure to be no better than either albumin or total protein in determining which patients will benefit from surgery.

Another way to make such an "across-the-board" comparison is via the contrast

$$H1: \frac{1}{13}\sum_{s=1}^{13}\{(1-q_s) - \frac{1}{2}[(1-G_2F_2^{-1}(p_s)) + (1-G_3F_3^{-1}(p_s))]\},$$

i.e., at each specificity, the sensitivity of the test based upon the K-G score is compared to the average performance of the tests based upon the continuous variables.

Alternatively, we might consider comparing the K-G based test against each of the continuous separator based tests separately via

$$H2: \frac{1}{13}\sum_{s=1}^{13}\{(1-q_s) - (1-G_2F_2^{-1}(p_s))\}$$

and

$$H3: \frac{1}{13}\sum_{s=1}^{13}\{(1-q_s) - (1-G_3F_3^{-1}(p_s))\}.$$

Such linear combinations are readily examined given the approximate multivariate normality of the vector of estimated differences. Table 14

Table 14. Estimated variance/covariance matrix of estimated differences in Sensitivities

Columns 1-13:

0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0005	0.0016	0.0014	0.0014	0.0026	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0016	0.0104	0.0095	0.0052	0.0089	0.0004	0.0001	0.0001	0.0001	0.0001	0.0001
0.0000	0.0000	0.0014	0.0095	0.0154	0.0108	0.0085	0.0015	0.0003	0.0005	0.0005	0.0005	0.0005
0.0000	0.0000	0.0014	0.0052	0.0108	0.0136	0.0101	0.0021	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0026	0.0089	0.0085	0.0101	0.0644	0.0414	0.0009	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0004	0.0015	0.0021	0.0414	0.0995	0.0024	0.0006	0.0006	0.0006	0.0006
0.0000	0.0000	0.0000	0.0001	0.0003	0.0000	0.0009	0.0024	0.0061	0.0039	0.0039	0.0039	0.0039
0.0000	0.0000	0.0000	0.0001	0.0005	0.0000	-.0003	0.0006	0.0039	0.0036	0.0036	0.0036	0.0036
0.0000	0.0000	0.0000	0.0001	0.0005	0.0000	-.0003	0.0006	0.0039	0.0036	0.0036	0.0036	0.0036
0.0000	0.0000	0.0000	0.0001	0.0005	0.0000	-.0003	0.0006	0.0039	0.0036	0.0036	0.0036	0.0036
0.0000	0.0000	0.0000	0.0001	0.0005	0.0000	-.0003	0.0006	0.0039	0.0036	0.0036	0.0036	0.0036
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0010	0.0062	0.0057	0.0028	0.0052	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0026	0.0124	0.0173	0.0135	0.0167	0.0033	0.0006	0.0012	0.0012	0.0012	0.0012
0.0000	0.0000	0.0031	0.0102	0.0153	0.0187	0.0211	0.0039	0.0003	0.0007	0.0007	0.0007	0.0007
0.0000	0.0000	0.0016	0.0055	0.0044	0.0073	0.0265	0.0067	0.0005	-.0001	-.0001	-.0001	-.0001
0.0000	0.0000	0.0000	0.0002	0.0007	0.0021	0.0059	0.0187	0.0017	0.0004	0.0004	0.0004	0.0004
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0031	0.0031	0.0021	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0001	0.0003	0.0000	0.0019	0.0013	-.0001	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0001	0.0003	0.0000	0.0019	0.0013	-.0001	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0001	0.0003	0.0000	0.0019	0.0013	-.0001	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0001	0.0003	0.0000	0.0019	0.0013	-.0001	-.0003	-.0003	-.0003	-.0003

Columns 14-26:

0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0010	0.0026	0.0031	0.0016	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0062	0.0124	0.0102	0.0055	0.0002	0.0000	0.0001	0.0001	0.0001	0.0001
0.0000	0.0000	0.0000	0.0057	0.0173	0.0153	0.0044	0.0007	0.0000	0.0003	0.0003	0.0003	0.0003
0.0000	0.0000	0.0000	0.0028	0.0135	0.0187	0.0073	0.0021	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0052	0.0167	0.0211	0.0265	0.0059	0.0031	0.0019	0.0019	0.0019	0.0019
0.0000	0.0000	0.0000	0.0000	0.0033	0.0039	0.0067	0.0187	0.0031	0.0013	0.0013	0.0013	0.0013
0.0000	0.0000	0.0000	0.0000	0.0006	0.0003	0.0005	0.0017	0.0021	-.0001	-.0001	-.0001	-.0001
0.0000	0.0000	0.0000	0.0000	0.0012	0.0007	-.0001	0.0004	0.0000	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0000	0.0012	0.0007	-.0001	0.0004	0.0000	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0000	0.0012	0.0007	-.0001	0.0004	0.0000	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0000	0.0012	0.0007	-.0001	0.0004	0.0000	-.0003	-.0003	-.0003	-.0003
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0042	0.0078	0.0062	0.0031	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0078	0.0327	0.0302	0.0031	-.0014	0.0000	0.0006	0.0006	0.0006	0.0006
0.0000	0.0000	0.0000	0.0062	0.0302	0.0481	0.0155	0.0040	0.0000	0.0003	0.0003	0.0003	0.0003
0.0000	0.0000	0.0000	0.0031	0.0031	0.0155	0.0348	0.0131	0.0031	0.0026	0.0026	0.0026	0.0026
0.0000	0.0000	0.0000	0.0000	-.0014	0.0040	0.0131	0.0153	0.0031	0.0017	0.0017	0.0017	0.0017
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0031	0.0031	0.0042	0.0021	0.0021	0.0021	0.0021
0.0000	0.0000	0.0000	0.0000	0.0006	0.0003	0.0026	0.0017	0.0021	0.0019	0.0019	0.0019	0.0019
0.0000	0.0000	0.0000	0.0000	0.0006	0.0003	0.0026	0.0017	0.0021	0.0019	0.0019	0.0019	0.0019
0.0000	0.0000	0.0000	0.0000	0.0006	0.0003	0.0026	0.0017	0.0021	0.0019	0.0019	0.0019	0.0019
0.0000	0.0000	0.0000	0.0000	0.0006	0.0003	0.0026	0.0017	0.0021	0.0019	0.0019	0.0019	0.0019

presents the estimated variance/covariance matrix of these differences (the values of these differences are reported in Table 13).

Table 15 below reports estimates of each of these contrasts, their estimated standard errors, the resultant value of the test statistic, and the associated two-sided p-value computed from the Standard Normal distribution. At a 10% level, none of the contrasts are found to differ significantly from zero. This suggests that there is no difference in overall diagnostic abilities of the three procedures. However, the lack of a statistically significant difference might be due to low power owing to the small sample sizes employed here.

Table 15. Estimates of contrasts

Contrast estimate	SE estimate	Contrast/SE	p-value (two-sided)
H1: -0.0223	0.0466	-0.48	0.62
H2: -0.0645	0.0557	-1.16	0.25
H3: 0.0199	0.0489	0.41	0.69

One of the reasons for the popularity of the ROC curve is that its area provides a single measure of the performance of the diagnostic test. This dissertation has argued that application of standard ROC-curve comparison methods might not be desirable in certain situations and has presented an alternative method for the comparison of tests when one is based upon a discrete separator. In conjunction with the method-

ology described herein, the use of contrasts can be seen to allow also the comparison of single summary measures of overall test performance. In fact, the contrasts H2 and H3 are closely allied to the comparison of the areas under ROC curves. For, suppose a test has the K (1-Specificity) values $\{\alpha_1,\alpha_2,\ldots,\alpha_K\}$ and associated sensitivities $\{\beta_1,\beta_2,\ldots,\beta_K\}$. Connecting each of the pairs $\{\alpha_i,\beta_i\}$, $i = 1,\ldots,K$, with line segments describes what might be interpreted as the ROC "curve," (ROC "polygon" might be preferable) for the test having a discrete separator variable. The area under the "curve" is then

$$\frac{1}{2}\left[\sum_{i=1}^{K-1}\beta_i(\alpha_{i+1}-\alpha_{i-1}) + \beta_K(\alpha_K-\alpha_{K-1})\right]$$

where, for convenience, we define $\alpha_0 = 0$.

When comparisons of such areas for tests having common specificities are made, the $(\alpha_{i+1}-\alpha_{i-1})$ $i = 1,2,\ldots,K-1$ and $(\alpha_K-\alpha_{K-1})$ are identical for each test so that the difference in areas under the curves is a linear function of the differences in sensitivities between the tests. H2 and H3 are of such a form, and so are similar to the comparisons of areas under ROC curves.

5. BIBLIOGRAPHY

Ahlquist, D. A., D. B. McGill, S. Schwartz, and W. F. Taylor. 1985. Fecal blood levels in health and disease: A study using HemoQuant. The New England Journal of Medicine 314(May 30):1422.

Anderson, T. W. 1984. An introduction to multivariate statistical analysis. 2nd ed. John Wiley and Sons, Inc., New York.

Bamber, D. 1975. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. J. Math. Psychol. 12:387.

Begg, C. B. 1986. Statistical methods in medical diagnosis. Pages 1-22 in the CRC critical reviews in medical information. Vol. 1, Issue 1. CRC Press, Inc., Boca Raton.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland. 1975. Discrete multivariate analysis: Theory and practice. The MIT Press, Cambridge.

Chung, K. L. 1974. A course in probability theory. 2nd ed. Academic Press, Inc., Orlando.

David, H. A. 1981. Order statistics. 2nd ed. John Wiley and Sons, Inc., New York.

Delong, E. R., D. M. Delong, and Daniel L. Clarke-Pearson. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 44:837-845.

Dorfman, D. D., and E. Alf. 1969. Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals: Rating method data. J. Math. Psychol. 6:487.

Green, D. M., and J. A. Swets. 1966. Signal detection theory and psychophysics. John Wiley and Sons, Inc., New York.

Greenhouse, S. W., and N. Mantel. 1950. The evaluation of diagnostic tests. Biometrics 6:399.

Hanley, J. A., and B. J. McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29.

Hanley, J. A., and B. J. McNeil. 1983. Method for comparing the area under the ROC curves derived from the same cases. Radiology 148:839.

Linnet, K. 1987. Comparison of quantitative diagnostic tests: Type I error, power, and sample size. Statistics in Medicine 6:147.

Randles, R. H., and D. A. Wolfe. 1979. Introduction to the theory of nonparametric statistics. John Wiley and Sons, Inc., New York.

Silverberg, E., and J. A. Lubera. 1983. A review of American Cancer Society estimates of cancer cases and deaths. CA 33:2.

Weinstein, M. C., and H. V. Fineberg. 1980. Clinical decision analysis. W. B. Saunders, Philadelphia.

White, A. A., and J. R. Landis. 1982. A general categorical data methodology for evaluating medical diagnostic tests. Commun. Statist.-Theor. Meth. 11:567.

Wieand, S., M. Gail, K. James, and B. James. 1987. A family of non-parametric statistics for comparing diagnostic tests with paired or unpaired data. To appear in Biometrika.

Yerushalmy, J. 1947. Statistical problems in assessing methods of medical diagnosis with special reference to X-ray techniques. Public Health Rep. 62:1432.

Youden, W. J. 1950. Index for rating diagnostic tests. Cancer 3:32.

6. ACKNOWLEDGMENTS

Uncountable thanks go to Dr. H. S. Wieand for his guidance, support and encouragement. Any shortcomings of my dissertation are my doing. Any "shiny" spots are largely due to him. Also, many thanks to Dr. H. T. David for his guidance and support and friendship.

Thanks again and again and again to my wife and family who, without hesitation, said 'sure' when I told them I wanted to return and get the Ph.D.

Finally, much appreciation, thanks and admiration go to Sharon Shepard for the outstanding job she did in typing this dissertation.